جامعة مؤتة
عمادة الدراسات العليا

بناء اختبار محكي المرجع لقياس كفايات المعلمين في بناء اختبارات التحصيل

إعداد الطالب
عيسى عبدالوهاب سليم الطراونة

إشراف
الدكتور عبدالله الصمادي

رسالة مقدمة إلى عمادة الدراسات العليـــــــــــا
استكمالاً لمتطلبات الحصول على درجـــة الماجستير
في القياس والتقويم قسم الارشاد والتربية الخاصة

جامعة مؤتة، 2006

بسم الله الرحمن الرحيم

MUTAH UNIVERSITY
Deanship of Graduate Studies

جامعة مؤتة
عمادة الدراسات العليا

نموذج رقم (14)

إجازة رسالة جامعية

تقرر إجازة الرسالة المقدمة من الطالب عيسى عبدالوهاب الطراونة الموسومة بـــ:

تطوير اختبار محكي المرجع لغايات المعلمين في بناء اختبارات التحصيل

استكمالا لمتطلبات الحصول على درجة الماجستير في القياس والتقويم.

القسم: الإرشاد والتربية الخاصة.

	التوقيع	التاريخ	
د. عبدالله عبدالغفور الصمادي		2006/5/11	مشرفاً ورئيسا
د. ساري سليم سواقد		2006/5/11	عضواً
د. عبدالله الرواجفة		2006/5/11	عضواً
د. أحمد عبدالحليم عربيات		2006/5/11	عضواً

عميد الدراسات العليا

أ.د. أحمد القطامين

MUTAH-KARAK-JORDAN
Postal Code: 61710
TEL :03/2372380-99
Ext. 5328-5330
FAX:03/ 2375694
e-mail: dgs@mutah.edu.jo sedgs@mutah.edu.jo
http://www.mutah.edu.jo/gradest/derasat.htm

مؤته – الكرك – الاردن
الرمز البريدي :61710
تلفون :03/2372380-99
فرعي 5328-5330
فاكس 03/2 375694
البريد الالكتروني
الصفحة الالكترونية

شكر وتقدير

الشكر كل الشكر، والحمد كل الحمد، لله عز وجل، الذي منحني القدرة على إتمام هذه الدراسة وتجاوز كافة صعوبتها. والصلاة والسلام على أسوتنا وقدوتنا سيدنا محمد صلى الله عليه وسلم وبعد:

فلا يسعني بعد أن شارفت هذه الرسالة على الانتهاء، إلا أن أقدم خالص شكري وتقديري إلى الدكتور الفاضل عبدالله الصمادي، الذي تفضل بالإشراف على هذه الرسالة منذ أن كانت مجرد فكرة تجول بخاطري إلى أن أصبحت حقيقة وواقعا، حيث كان الأنموذج والقدوة والعطاء، ولم يبخل في تقديم النصح والإرشاد والتوجيه الدائم والأفكار القيمة كلما دعى الأمر إلى ذلك. مما كان له كبير الأثر في إبراز هذا الجهد المتواضع إلى عالم العلم والمعرفة.

والشكر موصول لمن تفضل بقبول مناقشة هذه الرسالة وهم الدكتور ساري سواقد والدكتور احمد عربيات والدكتور عبدالله الرواجفه، وشكر خاص الى الزملاء معاذ المجالي وابراهيم الصرايرة على مساعدتهم لي وتشجيعهم ، كما اتقدم بالشكر الى الزميل محمد الطراونه الذي قام بتدقيق هذه الرسالة لغوياً .

والله ولي التوفيق

عيسى الطراونه

فهرس المحتويات

قائمة الجداول

قائمة الأشكال

قائمة الملاحق

الملخص

بناء اختبارمحكي المرجع لقياس كفايات المعلمين في بناء اختبارات التحصيل

عيسى عبدالوهاب الطراونه

جامعة مؤتة، 2006

هدفت هذه الدراسة إلى بناء اختبار محكي المرجع لقياس كفايات المعلمين في بناء اختبارات التحصيل. وتكون الاختبار من (70) فقرة توزعت على ستة مجالات تغطي ما يحتاجه المعلم من معارف ومهارات عند بنائه لاختبارات التحصيل وهي: الأهداف التعليمية، جدول المواصفات، أنواع الفقرات الاختبارية وصياغتها، انتقاء الفقرات وتهيئة الاختبار بشكله النهائي، خصائص الاختبار الجيد، تصحيح الاختبار وتفسير نتائجه.

وتم تحليل فقرات الاختبار حسب الطريقة التي تعتمد على قياس مجموعتين مختلفتين من الأفراد في وقت واحد، إحداها تلقت التدريب فيما له علاقة بمحتوى الاختبار، وعددها (40) والأخرى لم تتلق التدريب وعددها (40)، وتم استخراج معامل الصعوبة والتمييز لكل فقرة ووجد أن غالبية الفقرات تقع في مدى صعوبة يصل من (0.3) إلى (0.75). وتم التوصل إلى مؤشرات ثبات الاختبار الكلي بطريقتين: التجزئة النصفية وبلغت قيمته (0.88)، والاتساق الداخلي باستخدام معادلة كودر ريتشاردسون (Kr-20) وقيمته (0.88).

وبالنسبة لصدق الاختبار، فقد توافرت له ثلاثة أنواع من الصدق هي: صدق المحتوى، والصدق التمييزي، صدق البناء العاملي، وللتعرف على مستوى أداء معلمي المرحلة الثانوية، طبق الاختبار في منتصف العام الدراسي 2006/2005 على عينة الدراسة التي تكونت من (637) معلما ومعلمة من معلمي المرحلة الثانوية في كل من مديرية الكرك والزرقاء وجرش، اختيروا بالعشوائية العنقودية متعددة المراحل، وأظهرت النتائج أن معرفة معلمي المرحلة الثانوية لم تصل إلى المستوى المقبول وهو (64%).

Abstract

Constructing a criterion-referenced test to measure teachers' competencies in constructing achievement tests

Esa Abdelwahab Al-tarawneh

Mu'tah University

This study aimed at constructing a criterion –referenced test to measure teachers' competencies in constructing achievement tests. The test consisted of 70 items covered six different domains of test construction: educational goals, table of specifications, item format, item selection, for psychometric characteristics, and scoring and interpretation of individual performance.

A group of 40 trained teachers and another equal untrained group were utilized to derive the psychometric characteristics of the test. Item difficulties ranged between (0.3) and (0.75). The cut of percent score (64%)was defined as a criterion of master / non-master level baseline.

Both Internal Consistency and Split-half reliability indices were 0.88. The test has also acceptable properties of construct, discriminative, and content validity.

A sample of 637 male and female teachers was selected from Karak, Zarqa, and Jarash provinces during 2005/2006 schoolastic years. Results showed the mean performance was less than the criterion, which means that there is a clear weakness in test construction related competencies.

الفصل الأول

خلفية الدراسة وأهميتها

1.1 المقدمة:

تلعب الاختبارات بشكل عام واختبارات التحصيل بشكل خاص دوراً هاماً في العملية التعليمية-التعلمية (Teaching-Learning Process)، وهي تؤدي إلى تحسين العملية ومخرجاتها، من خلال العديد من القرارات التي يمكن اتخاذها في ضوء نتائج الاختبارات عبر المراحل المختلفة للعملية، وتعتمد صحة القرارات المتخذة على نوع ودقة المعلومات والنتائج التي يقدمها الاختبار.

لذا فقد نالت حركة الاختبارات اهتماماً بالغاً لدى المنشغلين في الميدان التربوي، من اجل بناء اختبارات جيدة تقدم نتائج ومعلومات موضوعية وواقعية يمكن اعتمادها في صنع القرارات ورسم السياسات التربوية (سواقد، 1987). فالاختبارات التحصيلية تقوي العملية التدريسية وتدعمها، فهي- إذا أحسن بناؤها- تكون عونا للمعلم والطالب على حدٍ سواء، فالمعلم يستطيع أن يوظّفها في تقويم الاستعداد للتعلم، ومراقبة تقدم العملية التعليمية، وتشخيص صعوبات التعلم وتقويم نتائجه. أمّا بالنسبة للطالب فهي تعد وسيلة مهمة في يده، فبناءاً على نتائجه في الاختبارات يتم تحديد نشاطه الدراسي، وأسلوب قراءته فهي تثير دافعيته للدراسة وهمته فيها.

ويلجأ المعلمون لهذه الاختبارات لحاجتهم إلى وسائل يقيسون بواسطتها مدى التغير الذي يحدث لدى طلابهم نتيجة لعملية التعلم. وعلى الرغم من المناداة بالبحث عن او تفعيل وسائل اخرى لتقييم الطلبة، إلاّ أن الاختبارات مازالت الوسيلة الاكثر قبولاً وشيوعاً. وعليه فإن إعداد الاختبارات المدرسية بشكل صحيح يشكل مهمة اساسية للمعلمين يكتشفون من خلالها مدى التغير والتعلم الذي احدثته عملية التعلم.

ولقد أشارت نتائج دراسات عربية إلى أن الاختبارات المدرسية التي يضعها المعلمون تشكل اهم مصادر المعلومات المستخدمة لاتخاذ مختلف القرارات ذات العلاقة بالطالب كما أشار خبراء الامتحانات في البلاد العربية في الاجتماع المنعقد

في الكويت عام 1974 إلى أن الاختبارات المدرسية هي المعيار الوحيد لتصنيف قدرات الطلبة والمفاضلة بينهم (المنظمة العربية للتربية والثقافة والعلوم، 1974،ص 11).

ولعل أبرز ما يدعو ويعزز استخدام الاختبارات التحصيلية هو التوسع في التعليم، فلم يعد التعليم مقصوراً على القلة اللذين يستطيع المعلم مراقبتهم عن قرب، ويتحقق من تغير سلوكهم، بل اصبح عدد المتعلمين كبيراً بحيث لا يستطيع المعلم مهما تفانى في اختيار أساليب التعليم من أن يتحقق بشكل كاف من حدوث التعلم على الوجه المطلوب، لذا استخدمت الإختبارات لتظهر مدى تقدم أو تأخر التلاميذ في تحصيلهم، ولتعطي صورة للمعلم عن مدى التغيرات السلوكية نتيجة التعلم، ولتحديد مواطن الضعف والقوة في عملية التعلم، وكذلك للمساعدة في تصنيف الطلاب حسب درجاتهم.

ومما لاشك فيه أن قياس وتقويم انجازات المتعلمين يثري تعلمهم. فعملية التعليم لا تنفصل عن عملية التقويم بل هما عمليتان متفاعلتان ومتكاملتان. والتقويم الجيد يعد شرطاً اساسياً للتعليم الفاعل. لذلك لابد أن يكون المعلم متمكناً من أساسيات القياس والتقويم التربوي والنفسي، وقادراً على تصميم وبناء الاختبارات التحصيلية الصفية (علام، 1995)، لأنها من أهم وسائل التقويم التي يقوم المعلم بها. فالمعلم أثناء عملية التدريس يقوم ببناء كثير من الاختبارات لطلابه، فعند بداية عملية التدريس يقوم المعلم بعمل اختبار قبلي،فقد يكشف هذا الاختبار عن عدم إتقان الطلاب للاهداف السابقة مما يستدعي إعادة تدريسها.واثناء عملية التدريس لابد من أن يتأكد من سلامة سير العملية التدريسية عن طريق عمل تقويم تكويني(بنائي) وهذا النوع من التقويم يسير جنباً الى جنب مع عملية التدريس لكي يساعد المعلم على معرفة مستوى الطلبة ومدى تمكنهم أو إتقانهم لها، ويزود الطالب بالتغذية الراجعة، ويساعد في تحديد اخطائه. والاختبار الجيد يعتبر موقفاً تعليماً الى جانب كونه أداة تقييم.

وفي نهاية عملية التدريس يأتي دور التقويم الختامي الذي يتم فيه عمل اختبارات تحصيلية تقيس ما تم التوصل اليه، وهذه الاختبارات في جميع مراحلها يعدها المعلم

وهي تقرر مصير الطالب وتحدد نجاحه أو فشله. وهذا الدور الكبير للاختبارات في عملية التعليم يتطلب أن يكون المعلم على معرفة بالشروط التي يجب مراعاتها عند إعداد الاختبارات، لضمان سلامة بنائها وامكانية تطبيقها ووضوح كيفية تفسير نتائجها، ومن هنا فإن الضرورة تقتضي ان يكون المعلم مكتسباً لمجموعة المعارف والمهارات اللازمة لبناء اختبارات التحصيل الجيدة، وهذه لا يمكن اكتسابها إلا بوجود برامج معدة إعداداً جيداً خلال عملية التأهيل داخل المؤسسات التعليمية أو التدريب أثناء الخدمة. ويذكر كل من ستجن وكونكلن(Stiggin & Conklin,1988) إنّ جزءاً كبيراً من برامج إعداد المعلم في مجال القياس والتقويم يجب أن تكرّس لتدريبه على الاختبارات الصفية. وحتى يقوم أي برنامج تدريبي ينبغي أن يبنى على المعرفة لحاجات المتدربين، لكي يؤدي دوره بالشكل المطلوب، وتقل احتمالات فشله، وبالتالي ينعكس إيجابا على التدريب والمتدرب بما لا يؤدي إلى هدر الأموال والجهود (أبو زيد ،1996).كما أن اي برنامج تدريبي ينبغي أن ينطلق من حاجات المتدربين في هذا المجال. وعليه فإننا إذا ما أريد اعداد برنامج تدريبي للمعلمين على بناء الاختبارات التحصيلية، فإن علينا معرفة حاجاتهم من التدريب في هذا المجال،وبما أن معظم الدراسات التي تناولت حاجات المعلمين في مجال بناء الاختبارات التحصيلية كانت تقوم على ما يلي :

أ. تحليل الإختبارات التي وضعها المعلمون، وهذا يمثل البعد التطبيقي الذي لا يكفي للوقوف على كافة احتياجات المعلمين في مجال بناء اختبارات التحصيل .

ب. أسلوب الإستبانة بغرض قياس كفايات المعلمين في مجال بناء اختبارات التحصيل، ويعد من أساليب التقرير الذاتي للمستجيب والذي وجه إليه كثير من النقد في الأدب التربوي. فهو قد يترك فرصة للمستجيب كي يزيف في إجابته حتى ليظهر بالمظهر المرغوب اجتماعياً، وحتى لا يظهر عدم معرفته بالمعلومات التي يسأل عنها.

وحيث أنه في الاختبار محكي المرجع يكون تفسير أداء الفرد بدلالة مجموعة كفايات أو أهداف معرفة جيداً دون الاشارة الى مستوى اداء الأفراد الأخرين من

المجموعة، حيث يحدد مستوى الاداء المقبول مسبقاً، وبالتالي يعطي هذا المحك معياراً مطلقاً للحكم على ما إذا كان الفرد قد اتقن الاداء أو لم يتقنه ،كما يمكن أن يكشف عن مواطن القوة والصعف في تحقيق هذه الاهداف.

ومن هنا تأتي هذه الدراسة بغرض بناء اختبار محكي المرجع لقياس المعارف والكفايات اللازمة لبناء اختبارات التحصيل لدى المعلمين، وتحديد مستوى درجة تمكن المعلمين لتلك المعارف والكفايات، من خلال أداة تعكس ذلك المستوى.

2.1 مشكلة الدراسة:

إن الحاجة الملحة إلى بناء أداة لقياس المعارف والكفايات اللازمة لبناء اختبارات التحصيل لدى المعلمين، من أجل المساعدة في الكشف عن مستوى المهارة لدى المعلمين في بناء هذا النوع من الاختبارات، وتحديد أهم الكفايات التي وصل فيها المعلمون إلى مستوى عالٍ أو مستوى متدن، ونقص مثل هذه الادوات، أظهرت مشكلة الدراسة التي هدفت الى بناء اختبار محكي المرجع لقياس كفايات المعلمين في بناء اختبارات التحصيل، ويمكن لهذه الدراسة تحقيق هدفها إذا أجابت عن الاسئلة التالية:

1) ما هي الخصائص السيكومترية لهذا الاختبار؟

2) ما مستوى أداء المعلمين على هذا الاختبار وهل يختلف هذا المستوى عن المستوى المقبول؟

3.1 أهمية الدراسة.

تأتى أهمية هذه الدراسة من أهمية الاختبارات كاداة للتقويم فهي ما زالت اداة التقويم الأكثر قبولاً وثقةً للحكم على أداء الأفراد والمقارنة بينهم أو مقارنتهم بمحك محدد. وما زال الكثير من قراراتنا يعتمد على نتائج تلك الاختبارات، وأن دقة تلك القرارات تعتمد إلى حد كبير على صحة بناء الاختبارات وقدرتها على تحديد مقدار السمة المقصودة، وهذا بدوره يعتمد على قدرة المعلم وكفايته في بناء الاختبار.

وهكذا فإن الاطمئنان على صلاحية الاختبارات وتوظيفها في اتخاذ القرارات يستوجب الاطمئنان إلى قدرة المعلم على بناءها، وهذا يتطلب وجود أدوات مناسبة.

وتأتي أهمية هذه الدراسة أيضاً من أنها تساعد المعلمين عامة على تحسين قدرتهم في تقويم وقياس السمات المطلوب تحقيقها لدى الطلبة، إذ أنها تضع بين أيديهم الكفايات الأساسية في بناء اختبارات التحصيل، كما أنها قد تساعد القائمين على برامج إعداد المعلمين، وتدريبهم وتأهيلهم لتطوير برامجهم نحو الأفضل بالكشف عن أهم جوانب القوة والضعف لدى المعلمين فيما يتعلق ببناء الإختبارات المدرسية، مما يؤدي بالتالي إلى تحسين وتطوير تعلم الطلبة.

4.1 التعريفات الإجرائية :

تكرر في هذه الدراسة مصطلح و كان لابد من تعريفه إجرائياً وهو:

الكفايات الأساسية لبناء الاختبارات التحصيلية: هي مجموعة المهارات التي ينبغي أن يمتلكها المعلم، كي يتمكن من خلالها ممارسة عملية بناء الاختبارات التحصيلية بشكل جيد، وتقاس من خلال الدرجة على كل فقرة وعلى كل مجال.

5.1 محددات الدراسة:

1) اقتصرت الدراسة على معلمي المرحلة الثانوية في الأردن، وبذلك يصعب تعميم نتائجها على معلمي المراحل الأخرى، بالرغم من إمكانية تشابه أو تماثل المعلمين في مختلف المراحل.

2) لم يكن من الممكن جمع المعلمين وتطبيق الاختبار عليهم في جلسة واحدة ولذلك كان التطبيق فردياً وكل معلم في موقعه.

الفصل الثاني

الإطار النظري والدراسات السابقة

1.2 الإطار النظري

1.1.2 مقدمة:

تشكل علوم القياس جانباً مهماً في دراسة الظواهر المختلفة وتقديرها. وكلما ازداد القياس دقة وموضوعية تحسنت درجة فهمنا للظاهرة، وأدى هذا إلى دقة في التنبؤ، وما يتبع ذلك من دقة في الضبط والتحكم.

وحيث أن الاختبار التحصيلي من عمل المدرس، هو الأداة الرئيسية لتقويم الطلاب في المدرسة، فإن الأداة المطلوبة هي اختبارات تتصف بالدقة والموضوعية بحيث يمكن بواسطتها تحديد مستويات الطلبة التعليمية بدرجة عالية من الدقة، وتمكن من اتخاذ القرار الملائم.

2.1.2 مفهوم القياس:

تعددت تعريفات القياس (أبو لبدة، 1985) فقد جاء :

تعريف كامبل(Campbell): القياس، هو تمثيل للصفات أو الخصائص بأرقام.

ويرى ستيفنز(Stevens) أن القياس في أوسع معانيه، هو عملية تحديد أرقام لأشياء أو أحداث وفقاً لقوانين، في حين عرف جلفورد (Guilford) القياس بانه وصف للبيانات أو المعطيات بالأرقام، وعرّفه نناللي (Nunnaly) على انه يتكون من قواعد استخدام الأعداد، بحيث تدل على الأشياء بطريقة تشير إلى كميات من الخاصية.

وعرّفه أوسترهوف(Oosstrhof,1994) على إنه العملية التي يتم إعطاء أرقام لخصائص الأشياء والأشخاص، بمعنى إننا لا نقيس الأشياء والأشخاص، بل نقيس خصائص الأشخاص أو الأشياء أو السمات، وبشكل أكثر دقةً فإننا نقيس السلوك الدال على وجود السمة.

ومع تعدد التعريفات الخاصة بمفهوم القياس، إلا أنها اتفقت على أن القياس عملية يتم بواسطتها التعبير عن خصائص الأشياء والأشخاص، أو مقارنتها بالاعتماد على رقم يدل على عدد من وحدات القياس التي يتم الاتفاق عليها.

3.1.2 أدوات القياس:

تعددت أدوات القياس التي من خلالها نحدد مواقع الأفراد على متصل السمة حسب درجة امتلاكهم لها. وقد صنف هوبكنز وانتس(Hopkin & Antes,1978) المشار إليها في (عودة، 1993) أدوات القياس وفق فئتين هما :

1) الأدوات التي تندرج تحت اسم الاختبارات وهي :

أ. الاختبارات المقننة أو المنشورة (Published Or Standardized).

ب. الاختبارات من إعداد المعلم (Teacher Made Test) .

2) الأدوات التي لا تندرج تحت اسم اختبارات، ويطلق عليها بشكل عام أدوات الملاحظة والتقدير ومنها: السجلات القصصية، وقوائم الشطب، وسلالم التقدير، ومقاييس العلاقات الاجتماعية والأدوات الاسقاطية.

4.1.2 الاختبارات:

تعد الاختبارات من اهم ادوات القياس النفسي والتربوي، وأكثرها شيوعاً في ميدان التربية وعلم النفس، وذلك لما تقدمه من معلومات واسعة ومؤشرات دالة ذات علاقة بالسمة المراد قياسها لدى الافراد.

ويؤكد مهرنز ولهمان(Mehrns and Lehmmn,1990) إن من أهم اهداف الاختبارات هو صنع القرارات التربوية، وتوفير بيانات يحرص واضعوها على تقديمها للمهتمين بالقضايا التربوية، وتصنف هذه الاختبارات وفقاً لتعدد أغراضها إلى التالي:

1) اختبارات الاستعداد العامة والمتعددة والخاصة(بطارية اختبارات).

2) اختبارات التحصيل (التشخيصية ذات الموضوع الواحد).

3) مقاييس الاهتمامات والشخصية والميول.

وصنفها غلاسر (Glaser,1963) المذكور في عودة(2002) حسب طريقة تفسير النتائج إلى قسمين هما :

أ)اختبارات معيارية المرجع(NRT)(Norm-Referenced-Test):

وفيه يتم تفسير أداء الممتحن عليها بالمقارنة مع متوسط أداء المجموعة التي ينتمي إليها وهذه المجموعة تعرف بالمجموعة المعيارية.

ب)محكية المرجع(CRT)(Criterion- Referenced- Test):

يتم تفسير أداء الفرد بالمقارنة مع مستوى أداء معين يتم تحديده مسبقاً من قبل واضع الاختبار ، دون الرجوع الى أداء الأخرين.

5.1.2 الاختبارات التحصيلية:

هي الاختبارات المستخدمة في المدارس لقياس تعلُّم الطلبة وتقويمهم، وهي في الأعم اختبارات من إعداد المعلمين أنفسهم، وهي منتشرة وشائعة،وقد برَّر كاري (Carey, 1994) استخدامها وشيوعها لعدة أسباب أهمها: أن المعلم يستطيع من خلالها قياس أنواع كثيرة من أنواع التعلم باختلاف مستوياته، وقدرتها على قياس مدى واسع من المحتوى، وسهولة إدارتها من حيث التطبيق، والتصحيح والتحليل بالنسبة لغيرها من الاختبارات. وقد كان لتوجه المعلمين وتركيزهم على قياس تحصيل طلبتهم وتقويمهم الأثر الأكبر في إعطاء هذه الاختبارات صبغة تحصيلية، حيث ركزت على مستوى تحصيل الطلاب للمعلومات والمهارات من المحتوى الدراسي، لذا فقد اشتهرت بأنها اختبارات تحصيلية، لتميزها عن غيرها من الاختبارات.

وقد عرَّف (عودة، 1993) الاختبار التحصيلي: بأنه طريقة منظمة لتحديد مستوى تحصيل الطالب لمعلومات ومهارات في مادة دراسية كان قد تعلمها مسبقاً بصبغة رسمية من خلال إجاباته على عينة من الأسئلة.

وتعد اختبارات التحصيل الصفية التي يجريها المعلم لطلابه نوعا من اختبارات القدرة، وهي تشكل الجزء الأهم من برنامج القياس والتقويم في المدرسة، كما يعد الاختبار التحصيلي أداة قياس يتم إعدادها وفق طريقةٍ منظمةٍ وقد ابرز كل

من جرينلند(Gronlund,1977) ،وهوبكنز(Hopkins,1978) الخطوات في إعداد الاختبار. ويمكن توضيح هذه الخطوات في الشكل [انظر الملحق(د)]

6.1.2 أغراض الاختبارات التحصيلية:

تتعدد أغراض الاختبارات التحصيلية تبعاً لمجالات استخدامها وزمان إجرائها، حيث يرى تايلر(Tyler,1988) أن لها ثلاثة أغراض وفقاً لوقت استخدامها في العملية التعليمية، فإذا استخدمت في بداية التدريس لتحديد من سيستفيد من برامج تربوية خاصة يكون غرضها الفرز او التصنيف، وإذا استخدمت خلال عملية التدريس فإن غرضها يكمن في التوصل إلى معلومات تساعد في مراقبة تقدم التلاميذ نحو النتاجات التعليمية المرغوبة وتنظيمها، الأمر الذي يستدعي بالضرورة تكييف التدريس لحاجات المتعلم، أما إذا استخدمت في نهاية عملية التدريس لقياس عملية التعلّم الذي حصل، فإن غرضها حينئذٍ يكون التقييم ومعرفة مواقع الطلبة على تدريج العلامة الكلية.

ويذكر (الزيود وعليان، 1990) مجموعة من الأغراض للاختبارات التحصيلية وهي :

1) التشخيص: إن هذه الاختبارات تساعد في تشخيص جوانب الضعف والقوة في تحصيل التلاميذ للمواد الدراسية، ويترتب على ذلك اتخاذ قرارات تربوية وإجراءات تهدف إلى تدعيم نواحي القوة وتعزيزها، والعمل على علاج الضعف وتلافيه.
2) التصنيف: في ضوء نتائج التلاميذ المتحققة على الاختبارات التحصيلية يمكن تصنيفهم إلى مستويات مختلفة، بهدف تنويع الأنشطة وأساليب التدريس بما يتلائم وتلك المستويات، ويلبي متطلبات التفاوت في جوانب المنهج المختلفة، وفي حاجات التلاميذ وميولهم وقدراتهم واستعداداتهم. وإلى جانب ذلك، فإن هذا التصنيف يساعد في توجيه التلاميذ إلى الفروع المختلفة.
3) ضبط التعلم وتوجيهه: تعدُّ الاختبارات المعدة إعداداً جيداً وسائلَ فعالةً في ضبط عملية التعلم وتوجيهها. ومن أبرز الفوائد في هذا المجال استثارة حوافز إيجابية

عند التلاميذ، وتوجيه نشاطهم نحو تحقيق أهداف مرغوب فيها، وإتاحة الفرصة أمام التلاميذ لتصويب الأخطاء التي يقعون فيها.

4) تثبيت التعلم: تتيح الاختبارات التحصيلية للتلاميذ فرصاً لممارسة نماذج من المعارف والمهارات التي سبق لهم تلقيها، وهذه النماذج تكون على شكل مهام أو أعمال وضعت بعناية في تلك الاختبارات، مما يجعل الاختبارات تؤدي غرضاً تعليمياً يتمثل في مساعدة التلاميذ على التعلم وعلى تثبيت ما تعلموه إلى جانب الغرض التقويمي الذي تضطلع به.

5) تحديد مسار حدوث التعلم: فالطريقة المتبعة في تقويم ما يتم تعلمه تساعد في تحديد مسار حدوث التعلم، نظراً لارتباط مجال التقويم بأنواع التعلم التي يهتم بها المنهج ومستوياته.

6) قياس مستوى التحصيل: ويقصد به تحديد مدى التقدم الذي أحرزه التلاميذ في اتجاه بلوغ الأهداف التربوية المقررة، أو مدى اكتسابهم للنتاجات التعليمية المراد منهم اكتسابها.

7) اتخاذ قرارات مناسبة تتعلق بعملية التطوير التربوي: إن نتائج اختبارات التحصيل تساعد صانعي القرارات التربوية في اتخاذ قرارات مناسبة حول عملية التطوير التربوي، فمن خلال ما يتوافر من معلومات عن مستوى الأداء الحالي وعن الظروف والإمكانات المتاحة في المدرسة والمعلومات ذات الصلة، يستطيع صانعو القرارات اتخاذ القرارات السليمة التي تؤدي إلى تحسين وتطوير العملية التعليمية.

7.1.2 الاختبارات المحكية المرجع والاختبارات المعياريّة المرجع:

من المعروف أنه عندما تزداد فاعلية البرنامج التعليمي. فإن درجات التلاميذ لا تتوزع وفقاً للمنحنى الإعتدالي، لأن كل تلميذ يظهر نواتج التعلم المتوقعة بمستوى مقبول، فيقل التباين بين درجات التلاميذ في الاختبار، وبالتالي يقل ثبات هذه الاختبارات، ويرجع انخفاض معامل الارتباط في الأساس إلى تقلص المدى بزيادة درجة التجانس بين الافراد، ولقد لفت جليزر(Glasser) نظر المربين وعلماء

القياس إلى عدم ملاءمة استخدام الأفكار التقليدية في القياس للمواقف التي تتضمن أساليب تعليمية جديدة، والى أهمية استخدام مدخل آخر لقياس نواتج التعلم أطلق عليه اسم (القياس محكي المرجع)(Criterion Referenced Measurement)، وعبر عن أفكاره واهتماماته في مقال نشر سنة 1963 م في إحدى الدوريات السيكولوجية بعنوان (تكنولوجية التعليم وقياس نواتج التعلم)، وفرق جلاسر في هذا المقال بين نوعين من القياس أحدهما القياس معياري المرجع، ويتركز الاهتمام فيه على ترتيب الفرد بالنسبة لأقرانه في القدرة التي يقيسها الاختبار، والأخر هو القياس محكي المرجع الذي يهدف إلى موازنة أداء الفرد بمستوى أداء مطلق، أو مستوى كفاية معين دون الحاجة إلى موازنة أدائه بأداء الأفراد الآخرين. ولقد أكد جلاسر على أهمية الوصف الجيد لمجال الأنماط السلوكية لدى المتعلم، ذلك ان درجة التلميذ على مقياس محكي المرجع توفر معلومات واضحة عما يستطيع عمله الفرد وعما لا يستطيع(جابر،1983).

وفي عام1969 بدأ بوبهام (Popham) عالم النفس الامريكي يتزعم حركة القياس المحكي الرجع . فقد نادى بالبدء بدراسات جدية لتحويل هذا النوع من القياس إلى واقع فعلي. ولذلك دعا إلى عقد مؤتمر متخصص عام 1970 لمناقشة القضايا والمشكلات السيكومترية المتعلقة بهذا المفهوم الجديد في القياس.

ونظراً للحداثة النسبية الاختبارات المحكية المرجع(Criterion Referenced Measurement) فقد اختلف علماء القياس أختلافاً ملحوظاً حول تعريف هذه الاختبارات تعريفاً دقيقاً. وربما يرجع سبب هذا الاختلاف الى غموض المقصود بكلمة محك. فقد أكد جلاسر أن المقصود بالمحك هو أن نوازن أداء الفرد بنطاق شامل من المعارف والمهارات المحددة تحديداً جيداً. وأكد أن الاختبار محكي المرجع هو ذلك الاختبار الذي يستخدم لتقدير أداء الفرد في نطاق من السلوك محدد تحديداً جيداً، أي أن عبارة " التحديد الجيد للنطاق السلوكي" هي الركيزة الأساسية لهذا المفهوم. ويعتمد تحديد المجال السلوكي على عدد الاهداف أو المهارات المطلوب قياسها. وهذا يتطلب صياغة الاهداف في عبارات سلوكية، أي في عبارات تصف التغير المطلوب احداثه في سلوك الطالب نتيجة ممارسته للخبرة التعليمية بحيث يمكن ملاحظة هذا

ملاحظة هذا السلوك وقياسه. وتختلف درجة هذا التحديد بحسب نوع ومحتوى النطاق السلوكي، ومدى تعقد الاهداف التي يشتمل عليها(علام،1986).

ويمكن القول أن الاطار المرجعي لتفسير نتائج الاختبارات محكية المرجع يركز على المهمات في ذاتها، ويستند الى معيار يمثل مستوى الاتقان المقبول. أما الاطار المرجعي لتفسير نتائج الاختبارات معيارية المرجع فيركز على الأداء النموذج لفئات من الناس، ويستند إلى معيار يمثل الأداء النموذج لفئة من الأفراد، أي التفسير بدلالة أداء المجموعة(ثورنديك وهيجن،1986).

وبينما نجد أن المقاييس معيارية المرجع قد صممت بحيث تحقق درجة معقولة من التباين في الاستجابة، وهكذا تمكن من المقارنة بين أداء الأفراد والجماعات المعيارية، نجد أن الاختبارات محكية المرجع ليس لها هذا المطلب، إن دور الاختبارات محكية المرجع هي أن تكون وصفية بقدر الإمكان، وليس لتباين الاستجابة إلا اهمية قليلة، وبسبب هذا التمييز، سوف يستطيع مطور الاختبار المرجعي المحك أن يحتفظ بعناصر اختبارية معينة أو يحتفظ بالاختبار ككل، في الوقت الذي لو التزم فيه بقواعد بناء الاختبار المرجعي المعيار فأنه يتخلص منه كله ،وعلى سبيل المثال، إذا كان أداء معظم المتعلمين جيداً في سلسلة من العناصر النتقاة من مجال سلوكي معين فإننا سوف نحتفظ بالعناصر حتى ولو كانت غير ذات اهمية لأغراض المقارنة التي تتطلبها الاختبارات معيارية المرجع، وذلك لأنها تحقق الأغراض الوصفية للاختبارات محكية المرجع(جابر،1983).

ويرى عودة (1993) أنه ليس من السهل أن نميز بين اختبارين إحداهما معياري والأخر محكي إلاّ إذا عرفنا جوانب الاختلاف، خاصة وأن أوجه الشبه ربما كانت أوضح من اوجه الاختلاف، ويمكن توضيح ذلك من خلال أوجه الشبه التالية :

1- ان كلاً منهما أعد لقياس أهداف تدريسية معينة، إلاّ أن بناء الاختبار المحكي ربما يحتاج إلى أهداف تفصيلية واكثر تحديداً. حيث يتطلب الاختبار المحكي درجة عالية من الانسجام بين الهدف والسؤال الذي يقيسه، فإذا كان الهدف التدريسي عاماً فهذا يعني احتمال صياغة عدة أسئلة تقع ضمن مجال الهدف.

2- أن كلاً منهما يقيس عينة من النواتج من خلال عينة من الفقرات، بمعنى أن كلاً منهما ينطبق عليه مفهوم العينة، ولكن نسبة حجم العينة إلى العدد الكلي المحتمل من الفقرات التي تغطي مجال المادة الدراسية تكون في الاختبار المحكي المرجع أعلى منها في الاختبار المعياري المرجع، وعلى الأغلب يكون المجال أضيق في المحكي المرجع.

3- بما أن الفرق الأساسي بينهما يكمن في طريقة تفسير النتائج، وأن أداء الطالب على الاختبار معياري المرجع يقارن بخصائص أداء طلاب صفه، فإن هذا الاختبار يؤدي الغرض كلما استطاع أن يميز بين أداء أي طالب عن الآخرين. بمعنى أن فقرات الاختبار لابد أن تتمتع بخصائص معينة بحيث تعطي أعلى تمايز، ويترتب على ذلك خصائص متعلقة بالثبات وانتقاء الفقرات بالتحليل الإحصائي. بمعنى أن خصائص الاختبار المتعلقة بالثبات، وطريقة إختيار الفقرات الداخلة في الاختبار لكل من النوعين سيكون مختلفاً نوعاً ما.

4- من المهم ان تعكس إجابة الطالب عن كل فقرة في اختبار محكي المرجع قدرته على الإجابة.

5- جميع الخصائص التي يتمتع بها الاختبار محكي المرجع، يمكن ان يتمتع بها الاختبار معياري المرجع ولكن بدرجات متفاوتة.

وقد حدّد الخطيب (1993) استعمالات الاختبارات محكية المرجع بما يلي :

1- إن استخدامها يساعد في اتخاذ القرارات لتقييم البرامج التعليمية فيما إذا كانت الإجراءات التعليمية ناجحة أم لا.

2- تساعد في تشخيص الصعوبات التعليمية وتخطيط التعليم.

3- تساعد في التوجيه والإرشاد لأعمال تتطلب مهارات خاصة.

4- تحديد درجة تمكن الطالب من اكتساب كفايات ومهارات معينة.

ويرى جرونلد Gronlund (1973) أن الفرق الاساسي بين اختبار المحك والاختبار المعياري يكمن في صعوبة أسئلة الاختبار، ففي الاختبار المعياري تصمم

الأسئلة عن قصد من أجل الحصول على توزيع أوسع للدرجات حتى يتسنى التوزيع الثابت للتلاميذ تبعاً لتحصيلهم. أما في اختبارات إتقان المحك فإن انتشار الدرجات أمر غير متوقع وغير مرغوب فيه، وصعوبة الاسئلة يجب أن تشتق بصورة مباشرة من عمليات التعلم المراد قياسها، فإذا لم يستطع أي من الافراد أو بعضهم الإجابة عن الاسئلة قبل الدراسة، ولكن استطاع كل الافراد الإجابة عليها عقب الدراسة، فإن هذا يدل على أن الأسئلة وأسلوب التدريس هي الأسئلة الفعالة والأسلوب الفعال. ويجب أن لا يفهم من ذلك أن صعوبة الأسئلة يمكن تجاهلها عند وضع اختبارات الإتقان للمحك. بل إنه على العكس يجب أن توجه عناية خاصة لمقارنة صعوبة الاختبار مع صعوبة عمليات الأداء التي تم وضعها في النتائج النوعية للتعلم، فالاختبار يجب أن يكون سهلاً وليس صعباً. ويتطلب الحصول على مستوى معين لصعوبة الأسئلة عناية فائقة أثناء كتابة الأسئلة.

والغرض من القياس المحكي المرجع هو اتخاذ قرار فيما إذا كان الفرد يمتلك المهارات والمعرفة للنجاح في المستوى اللاحق لأي سمة يتم قياسها أم لا. ويساعد كذلك في تحديد امتلاك الفرد للمهارات والمعرفة التي يتوقع أن يمتلكها في مرحلة زمنية معينة. والحقيقة أن القياس المعياري المرجع يعطي معلومات فقط بالنسبة إلى المجموعات المعيارية الامر الذي يعتبر قصوراً فيه. وتم تطوير القياس المحكي المرجع في محاولة للتغلب على هذا القصور. ويطور الاختبار المحكي المرجع لتحديد فيما إذا كان فرد معين يمكن أن يستجيب بشكل صحيح إلى فقرة أو مهمة خاصة. وكل طالب يقارن مع المهمة التي هو بحاجة إليها بواسطة فقرات الاختبار أكثر منه مع الأشخاص الذين أخذوا الاختبار. وقدرة الاختبار المحكي المرجع في التمييز بين الأفراد ليست مهمة، والمسألة المهمة هي تحديد فيما إذا كانت فقرة الاختبار تقدر الشيء الذي يريد المعلم تقديره، كذلك يمكن أن يكون مهماً لإعطاء معلومات حول المجال بشكل عام، وكل فقرة هي مهمة كل فقرة هي محك. وفي النظام المعياري المرجع من المهم أن ندرس الفقرات بمفردها، ولكن بشكل عام فإن الأهمية الكبيرة للاختبار الكلي، ليكون قادراً على التمييز في الأسلوب الذي

يريده المعلم. وبالنسبة للاختبارات المحكية المرجع، فإن المعايير للحكم على اختبار تتعلق مبدئياً للحكم على الفقرة بمفردها(Beggs and Lewis,1975).

ومن العيوب الأساسية لاستخدام الاختبارات المرجعية المعيار في التقويم التربوي، أنها لا تقيس الأهداف التعليمية التي يسعى المعلم إلى تحقيقها في الفصل الدراسي. ويعني ذلك وجود نقص وقصور بين ما يقيسه الاختبار وما يؤكده المنهج التعليمي في مدرسة معينة أو منطقة تعليمية معينة أو فصل معين، أي عدم اتساقها مع المنهج التعليمي. كما أنها لا تقدم الفائدة المرجوة في تشخيص مواطن القوة والضعف في تحصيل الطلاب حتى يستطيع المعلم أن ييسر وسائل المعالجة ورفع كفاءة العملية التعليمية. ومن عيوب هذه الاختبارات حين تستخدم لأغراض التقويم ما ينشأ عن حاجتها لإنتاج أو احداث درجة معقولة من التباين في الإستجابة. ويعني ذلك أن الفقرات الإختبارية التي لا تسهم في التباين بين المفحوصين تحذف عادةً في التجارب الاولى للاختبار، وعند تنقيحه، وسوف تتعرض الفقرات التي يجيب عنها معظم الأفراد في الإجابة عنها المفاهيم التي يعتبرها معظم المعلمين هامة. وهذا يعني أن اختبارات التحصيل المعيارية المرجع والمطوّرة تستبعد هذه العناصر لأنها لا تسهم في التباين في الاستجابة(علام، 1986 ؛ جابر، 1983).

إن المعلومات التي نحصل عليها من تفسير نتائج الاختبارات محكية المرجع أو اختبارات الكفاية تساعدنا في اتخاذ قرار فيما إذا كانت هناك حاجة إلى تدريب إضافي أم لا، في المهارات التي قاست مستوى كفاية الطالب فيها، وفيما إذا كانت البرامج التعليمية المستخدمة حالياً في المدرسة تعطي نتائج مرضية أم ان النتائج التي تعطيها ليست في المستوى المطلوب (الزيود وعليان ،1990).

وعند الحديث عن تحليل الفقرات في الاختبارات محكية المرجع مثل اختبارات التمكن، فإن متوسط معامل الصعوبة سيكون مرتفعاً، أي أن الفقرات سهلة نسبياً. ويعتمد إنتقاء الفقرات في الاختبارات محكية المرجع بشكل عام على درجة قياسها لنتائج التعلم المقصود بالدرجة الأولى، فقد تكون الفقرات صعبة جداً أو سهلة جداً ومع ذلك لا يتم تغييرها بقصد الحصول على قدرة تمييزة عالية، إلاّ أن هذا لايعني الإنقاص من أهمية الخاصية التمييزة للفقرة(عودة،1993).

ويرى عودة أن هناك الكثير من المهارات في جميع المستويات التعليمية لاتحتمل أن تفسر نتائج الاختبارات التي تقيس هذه المهارات على أنها معيارية المرجع، ولذلك بدأ الشعور بضعف دور الاختبارات معيارية المرجع في اغراض التقويم المختلفة مثل التشخيص،والاستعداد والاختيار،والتوجيه والارشاد.

8.1.2 تحليل نتائج الاختبارات المدرسية:

يعد تحليل نتائج الإختبارات المدرسية خطوة هامة وأساسية من خطوات التقويم بالإختبارات التي يعدها المعلم في العملية التدريسية، فعلى ضوء نتائج التحليل يتم اتخاذ قرارات تربوية متعددة، منها ما يتعلق بالطالب كفرد، أو بالطلبة في مجموعات، أو بالبرنامج التعليمي في مدخلاته، وخطوات تنفيذه ومخرجاته بما فيها الإختبار نفسه. وهذا ما أشار إليه عودة (1990) في دراسته التي هدفت إلى التعرف على مدى تطبيق المعلمين للتحليلات الإحصائية في اختباراتهم المدرسية، ومدى شعورهم بأهميتها بالنسبة لعملية التقويم، كما هدفت إلى التعرف على الأسباب التي قد تحول أو تعيق تطبيقها، وأثر بعض العوامل الشخصية والتربوية في تطبيقها أو الشعور بأهميتها. وأشارت الدراسة إلى ضعف واضح في هذا المجال، بمعنى أن الوصول إلى المستوى المطلوب ما زال بعيداً، وقد يحتاج ذلك دراسات وبحوث للتعرف إلى الأسباب ووسائل معالجتها.

ويشير عوده (1993) إلى أن المعلم هو المطور والمطبق والمصحح للاختبار ومفسر لنتائجه، وهذا يختلف عن الاختبارات المقننة التي تطور وتجهز لكل من يريد استخدامها. وفيما يتعلق بالتجريب الأولي للإختبار، فالمعلم لا يجرب اختباراته قبل تطبيقها، لأن الصف الذي يطبق عليه هو المجتمع وهو العينة، ولذلك يقدر المعلم زمن الإجابة، كما أن التعليمات ليست مبرمجة، والمجال مفتوح للأسئلة عن صياغة الفقرات، وأي تعليمات غامضة يمكن إعادة توضيحها من قبل المعلم.

أما بالنسبة للخصائص السيكومترية للاختبار، فهناك عدة مؤشرات للصدق والثبات وتختلف أهداف المؤشر بإختلاف الغرض من الاختبار، وبالنسبة للاختبارات من إعداد المعلم فإن أهم نوع من أنواع الصدق هو صدق المحتوى، ويحاول المعلم

توفير هذا النوع من الصدق من خلال تغطية الأهداف والمحتوى باستخدام لائحة المواصفات. أما بالنسبة لاشتقاق المعايير بغرض تفسير النتائج، فهي بالنسبة لامتحانات المعلم معايير محلية، وغالباً ما يحتفظ بها على شكل علامات خام دون تحويل، ويفسر النتائج بحساب بعض الإحصائيات مثل إحصائيات النزعة المركزية وإحصائيات التشتت.

وفيما يتعلق بصدق الاختبارات المدرسية، يرى هويست(1983) انه موضوع معقد مما قد يفسر عدم تقدم البحوث الملاحظ فيه. ومن المشكلات التي قد تعرقل تقدم الدراسات النظرية لصدق المحتوى عدم وجود طريقة مقبولة تمكن من قياسه.

2.2 الدراسات السابقة:

ظهرت العديد من الدراسات التي تناولت الاختبارات المدرسية التي يعدها المعلم، وبحثت هذه الدراسات في مجالات عديدة من مجالات التقويم، وسنحاول في هذه الدراسة استعراض الدراسات التي تعرضت لمهارات وممارسات المعلمين في تطوير الاختبارات التحصيلية.

إن أول الدراسات في موضوع الاختبارات التحصيلية في البلاد العربية قامت بها دائرة التربية في الجامعة الأمريكية في بيروت عام 1961، حيث ضمَّت في مؤتمر تربوي ممثلين عن عدد من البلاد العربية، وبعدها دراسة أجرتها الدائرة الثقافية في جامعة بيروت العربية عام 1964، وذلك في مؤتمر عقد في الجزائر، ضمَّ ممثلين لأكثر الدول العربية، تم فيه مسح شامل لأساليب الامتحانات المتبعة في تلك الدول، وفي كلا المؤتمرين جاءت توصيات واضحة بضرورة إعادة النظر في الإمتحانات في البلاد العربية، بقصد تحسينها وإرسائها على قواعد فنية مدروسة.

وفي حلقة توحيد أنظمة الاختبار والانتقال في المراحل الدراسية المختلفة التي عقدت عام 1970 في القاهرة، وفي اجتماع خبراء تطوير نظم الامتحانات في البلاد العربية الذي عقد في الكويت عام 1974 بدعوة من المنظمة العربية للتربية والثقافة والعلوم المشار إليه في(القرشي، 1986)، تم دراسة الاختبارات التي تجربها هذه الدول في مدارسها، وكان من أبرز ما أكدت عليه هذه الاجتماعات:

1- ضرورة تحديد الأهداف الدراسية لكل مادة وتطوير أساليب الامتحانات ووسائل التقويم المستخدمة.

2- إنشاء أجهزة فنية متخصصة في وزارات التربية تختص بنتائج الامتحانات وتقويمها والاستفادة منها.

3- إدخال مقرر دراسي في تقويم الاختبارات والمقاييس في مناهج معاهد وكليات المعلمين.

4- تدريب المعلمين على إعداد الاختبارات وطرق تقويم الطلبة.

5- إن الدول العربية ما زالت تعاني من مشكلات وتعقيدات ناتجة عن النظم التقليدية للامتحانات.

6- شيوع الاختبارات المقالية في قياس تحصيل الطلاب في معظم البلاد العربية.

7- ندرة وجود مراكز متخصصة للأبحاث في التقويم أو برامج منظمة لتدريب المعلمين على الأساليب الحديثة في التقويم.

ومن أولى الدراسات على المستوى العالمي والتي اهتمت في هذا الميدان، جاءت نتيجة للمؤتمر الدولي للإمتحانات الذي تبنته مؤسسة كرنجي في نيويورك المشار إليها في (عطية ، 1970) و اشتركت فيه لجان قومية مختلفة، وعلماء مثل سيرل بورت (Cyril Burt) وسبيرمان (Spearman) وبيرسون (Pierson)، وأبرزت هذه الدراسة عدة ملاحظات، أهمها:

1- قصور الأسئلة في الإحاطة بجوانب الموضوع الذي يجري فيه الإمتحان.

2- غياب التكافؤ في صعوبة الأسئلة من سنة إلى أخرى.

3- اختلاف المصحح الواحد مع نفسه من مرة إلى أخرى، وإختلاف تقدير المصححين بين بعضهم البعض.

ومن المعروف أن المعلمين يبنون اختباراتهم بأنفسهم، ولأغراض مختلفة، وعليه فانه لا بد من أن تكون لديهم المعرفة بطبيعة هذه الأغراض، وكيفية بناء الإختبارات، وقد أجرى كيربر (Kerber,1976) دراسة هدفت لتحديد مدى معرفة المعلمين لمفاهيم القياس والاختبارات، وطبقت هذه الدراسة على (328) معلماً و (43) طالباً في السنة الأخيرة من جامعة أيوا، حيث أستخدم فيها اختبار لقياس

المفاهيم الخاصة بالإختبارات، وقد توصلت إلى أن معرفة المعلمين تزداد بزيادة التدريب على استعمال الإختبارات، وبزيادة الخبرة في التعليم، ومستوى الدرجة العلمية، وأظهرت الدراسة عدم وجود فروق دالة إحصائياً بين علامات طلاب الجامعة، والمعلمين على اختبار مفاهيم القياس والاختبارات.

ومن الدراسات في هذا المجال، الدراسة المقارنة لواقع أساليب التقويم والامتحانات في دول الخليج العربي، التي أجراها كل من الصانع وشبلي ومعطي وقرشي ومكينزي (1981) المشار إليها في (القرشي، 1986)، حيث تم استطلاع آراء المسؤولين في إدارة الامتحانات بالإضافة إلى العاملين في الميدان التربوي من مدرسين و مديرين، وذلك للحصول على صورة حية لواقع التقويم التربوي، وقد شملت هذه الدراسة الجانب التنظيمي والإداري للامتحانات، كما تطرقت إلى البحوث والوسائل المتبعة في التقويم، كما أجرى الفريق نفسه دراسة ميدانية عام 1984 ، وكان من نتائج هاتين الدراستين:

1-إن غالبية الإدارات المسؤولة عن التقويم تنشغل بأمور روتينية متعلقة بتنظيم الامتحانات وإدارتها.

2- إن قلة معرفة ودراية المعلمين بالأساليب الحديثة في التقويم، قد جعلت عملية التقويم تقتصر في الغالب على الجانب المعرفي في مستوياته الدنيا.

3- ندرة المختصين في التقويم، وعدم توفر مراكز متخصصة أو أدوات ونماذج تقويم.

4-الاختبارات وهي الأسلوب السائد في تقويم الطلاب، وترتكز على الاختبارات المقالية.

ومن الدراسات التي اهتمت أيضاً بتحديد مهارة المعلمين في إعداد الاختبار دراسة قام بها نيومن (Newman,1981)، وكان هدف هذه الدراسة التعرف على فهم المعلمين لمهارات بناء الاختبارات، وبحث العلاقة بين كفاية الاختبارات الصفية وبعض المتغيرات كالمؤهل والخبرة، وقد استخدم فيها أداتين الأولى استبانة لتحديد فهم المعلمين لمبادئ الاختبار، والثانية استبانة أيضا لتقييم الاختبارات التي يعدها

المعلمون، وقد توصلت هذه الدراسة إلى أن لدى المعلمين فهما متوسطا لمبادئ الاختبار الصفي الرئيسية، وان كفاية المعلم تزداد مع الخبرة والمؤهل العلمي.

وفي دراسة أجراها علام(1985) والتي هدفت إلى بناء اختبار محكي المرجع لقياس تمكن المعلمين من المهارات الأساسية اللازمة لتطوير الاختبارات المدرسية، وتكون الاختبار من أربعة مجالات هي : صياغة الأهداف السلوكية، وتصميم خطة الاختبارات المدرسية، وبناء و تصحيح فقراتها، ثم تحليل البيانات المستخلصة من تطبيقها، وقد حدد النطاق السلوكي الذي يقيسه الاختبار، وصاغ الأهداف السلوكية، وتمت مراجعتها عن طريق خمسة من خبراء القياس، ثم حللت الأهداف النهائية، ثم رتبت في أطر تخطيطية هرمية، من حيث أولويتها في تحقيق الأهداف النهائية، وتم إعداد مجموعة مستقلة من مواصفات الاختبار خاصة بكل مهارة من المهارات الأساسية التي غطاها الاختبار بهدف تحديد المجال السلوكي الذي يقيسه كل اختبار فرعي تحديداً أكثر تفصيلاً، وتحقق من وضوح و اكتمال مواصفات الاختبار بواسطة خبراء في القياس، ثم تم بناء فقرتين متكافئتين لقياس كل هدف من أهداف المهارات الأربع، في ضوء مواصفات الاختبار التي أعدت في المرحلة الثانية، هذا وقد ارتبط مستوى صعوبة كل فقرة بمستوى الهدف الذي تقيسه ومستواه المعرفي، واختيرت للاختبار فقرات من نوع الاختيار من متعدد(3-5) بدائل، وفقرات أخرى تتطلب إجابة قصيرة، وهكذا فقد تم بناء اختبار مؤلف من (85) فقرة.

وأجرى جرادات (1988) دراسة هدفت للتعرف على مدى معرفة معلمي العلوم للمرحلة الاساسية في الأردن بكفايات بناء الاختبارات المدرسية وممارستهم لها، وأثر كل من الخبرة والدرجة العلمية، والجنس على هذه المعرفة، حيث طبق اختبار خاص بكفايات بناء الاختبارات المدرسية على (298) معلماً من معلمي العلوم للمرحلة الاساسية، وقام بتحليل إجاباتهم، كما عمل على تقدير درجة ممارستهم لبناء الاختبارات من خلال تحليله لـــ (120) اختبارِ أعدها المعلمون لطلبتهم، وذلك بالملاحظة المباشرة، وقد أشارت النتائج إلى أن معرفة المعلمين وممارستهم لهذه الكفايات لم تصل إلى المستوى المطلوب تربوياً، وأظهرت النتائج فروقاً دالةً إحصائياً لصالح حملة البكالوريوس فما فوق، ولذوي الخبرة القصيرة في

الميدان. وقد أوصى بزيادة الاهتمام بتأهيل المعلمين قبل الخدمة وأثناءها، لرفع درجة كفايتهم ومساهمتهم في بناء الإختبارات، وإجراء دراسات تشمل المرحلتين الاساسية والثانوية، وأخرى تعتمد الملاحظة المباشرة للاختبارات التي يجريها المعلمون لإختبار طلابهم، وتقويمها في ضوء معايير محددة وواضحة.

وقام مارسو وبيج (Marso and Pigge,1988) بتحليل للاختبارات من إعداد المعلم، من حيث إجراءات تطبيقها ، والمتطلبات المعرفية لإعدادها ، وأخطاء بناء الفقرات ، وشملت هذه الدراسة على تقارير المعلمين حول إجراءات الاختبار ، وعلى التقييم المباشر للاختبارات لعينة بلغت (175) إختبارا، والتي أعدها المعلمون . حيث كان التركيز على تقييم طبيعة ونوعية الاختبارات التي أعدها المعلمون في صفوف المدارس الحكومية ، ووصف هذه الاختبارات . وقد تبين أن مدرسي هذه الصفوف قد أعدوا وطبقوا العديد من الاختبارات الرسمية خلال العام الدراسي، واستخدموا بكثرة أسئلة الاختبار من متعدد ، وأسئلة المزاوجة، والإجابات القصيرة، ولكن نادرا ما يستخدمون أسئلة المقال ، أو يتموا التحليلات الإحصائية لنتائج اختباراتهم. وتبين كذلك أن معظم اختبارات المعلمين كانت فقراتها في مستوى المعرفة فقط وأن الفقرات تضمنت أخطاء في بنيتها بشكل واضح.

وحول بناء الاختبارات من إعداد المعلم قام جنتري (Gentry,1989) بتقديم ورقة في المؤتمر السنوي لجمعية الأبحاث لمنطقة الجنوب الوسطى التربوية في الولايات المتحدة الأمريكية، تقترح هذه الورقة ضرورة الفهم الواضح والمعرفة السليمة بالأهداف وجداول المواصفات، لأنها تؤدي إلى تطور بارز في نوعية الاختبارات التي يعدها المدرسون، وتتضمن كذلك إرشادات لبناء الاختبارات بأنواعها والصفات المشتركة في اختيار محتوى الاختبار، وتطوير جداول المواصفات وكذلك تم وصف الاختبارات المعيارية المرجع، والاختبارات المحكية المرجع بما فيها اختبارات الإتقان.

أما دراسة العمري (1991) فقد ركزت على معرفة الأساليب و الممارسات التقويمية لدى المعلمين في الأردن ،و لتحقيق ذلك وجهت الباحثة استفتاء للمعلمين حول مدى استخدامهم للأساليب التقويمية من خلال الاختبارات الشفوية و التحريرية

و أساليب التقارير أو المشاريع وأي أساليب أخرى ، وقد شملت الاستبانة ستة أبعاد هي : التخطيط للاختبار ،إخراجه،و تطبيقه،و تصحيحه،و تحسين فقراته ،ثم استخدام الاختبار في عملية التقييم ، وعرضت الاستبانة على المحكمين للحكم على مدى مناسبتها ، وحسب معامل الثبات باستخدام معادلة كرونباخ- الفا فوجد أنه (89%)،وقد تكونت العينة النهائية من(202)معلماً من معلمي المرحلة الأساسية.وأوضحت النتائج أن الاختبارات الكتابية هي الأكثر شيوعاً،حيث كانت نسبتها(86,6%)،وأن أكثر الفقرات الاختبارية شيوعاً هي فقرات المقال التي تتطلب إجابة قصيرة،إذ بلغت (58,4%)،و أن المعلمين قد أظهروا بعض الممارسات غير المقبولة تتعلق بإخراج الاختبار وتطبيقه كتصويب الأخطاء المطبعية أثناء أداء الطلبة للاختبار ووضع أسئلة قبيل موعد الاختبار بمدة غير كافية و إعطاء تعليمات شفوية أثناء أداء الطلبة للاختبار ،وأن المعلمين يضعون أسئلة لم يركزوا عليها أثناء تدريسهم .

وهدفت دراسة علام (1991)إلى إعداد اختبار تشخيصي مرجعي المحك في أساسيات القياس والتقويم التربوي و النفسي،وتكون الاختبار من خمسة مجالات،كان من ضمنها المجال الخاص ببناء الاختبارات الصفية وتقيميها،وقد حدد الباحث الأهداف السلوكية للمجالات الخمسة للاختبار،والتي رأى أهمية تعريف طلاب كليات التربية بها،وقد عرض هذه المجموعة من الأهداف على معلمين وموجهين ومشرفين على برامج التربية العلمية وعددهم(22)،بالإضافة إلى(7)من الأساتذة بغرض تقدير أولويات هذه الأهداف على ميزان ثلاثي التدريج وحذف الأهداف التي يقل متوسطهاعن(1,8)، وقد اشتمل الاختبار على(60)مفردة من نوع الاختيار من متعدد لقياس(60)هدفاً سلوكياً،بعد أن عرض الاختبار على سبعة من الأساتذة المتخصصين في المجال بهدف التحقق من أن كل مفردة تعد مقياساً صادقاً للهدف المناظر لها،وطبق الاختبار على(43) طالبا من طلاب الدبلوم العام في كلية التربية بجامعة الأزهر،وحسب معامل الثبات باستخدام معادلة كيودر(K-20)فوجد أنه (74%)،واستخرج معامل الاتفاق بطريقة سابكوفياك فأشير إلى أنه (0,68)،ولتحديد الحد الأدنى لمستوى الإنجاز في الاختبار استخدم الباحث طريقة ندلسكي وطريقة

إيبل وهما من الطرق المطلقة،كما استخدم طريقة المجموعات المتناقضة،و المجموعة الحدية من الطرق النسبية،وقد تراوحت درجة الحد الأدنى بين (40إلى44).

وكذلك قام وليامز (Williams,1991) بتقييم المعلمين النظاميين والخاصين من حيث إجادتهم للمهارات التي تحسن من نوعية الاختبار الذي يضعونه والذي يعتمد على المحتوى، وقدم خطوط عريضة لبناء الاختبارات الموضوعية والمقالية، ولكيفية انتقاء وتقييم الأسئلة مثل: الأسئلة القصيرة، والموضوعية، والمقالية، والإكمال، والصح والخطأ والمزاوجة، وأوصى باستخدام جداول المواصفات لتأكيد صدق محتوى الاختبار، ولتطوير أسئلة الاختيار من متعدد التي تقيس مهارات التفكير العليا، كما ويقترح البحث حلولاً للتخلص من الأخطاء البنائية والشكلية للاختبار، ويلقي الضوء على مساويء كل نوع من الأسئلة كما وقدم الباحث مبادئ لكتابة اختبار جيد على شكل ملاحق وهي : قائمة أفعال يستخدمها المعلم حسب تصنيف بلوم للأهداف المعرفية، ودليل للطلاب لفهم الكلمات في الأسئلة المقالية.

كما قدّم مارسو وبيج (Marso & Pigge,1992) تلخيصاً لنتائج مسح لـــ (225) دراسة في موضوع المعارف والمهارات لمعلمي الصفوف، والمتعلقة باستخدام وتطوير الاختبارات التي يضعونها، وتوحي النتائج إلى أن تغيراً قليلاً حدث في مهارات المعلمين الاختبارية في ربع القرن الأخير بعد أن كانت تنقصهم المعرفة في قياس وتقويم طلابهم وقلة الدورات التدريبية أثناء الخدمة، وتركيز اختباراتهم على مستوى التذكر من الأهداف، وعدم إتباعهم طرق تحسين الإختبارات مثل تحليل الفقرات، وتحتوي ورقة البحث هذه على جداول للممارسات والمواقف التي يتعرض لها المعلمون، ومعلومات ومهارات حددتها هذه الدراسة في مجال إعداد الإختبارات.

أما دراسة الحميضي(1993) والتي هدفت إلى قياس مدى معرفة وممارسة معلمات الدراسات الاجتماعية بالمرحلة الثانوية بمدينة الرياض في المملكة العربية السعودية بإعداد وتصحيح وتحليل الاختبارات، وتم بناء اختبار لقياس معرفة المعلمين، وحسب معامل ثبات الاختبار باستخدام طريقة الإعادة، ووجد أنه

(0.97)، وباستخدام طريقة التجزئة النصفية فكان(0.83)، كما أعدت استبانه حسب ثباتها بطريقة الإعادة فكان (0.93) وبطريقة التجزئة النصفية فوجد أنه (0.75)، وبطريقة كرونباخ-ألفا فبلغ (0.86). وتم تطبيق كل من الاختبار و الإستبانة على (171) معلمة اخترن عشوائياً. وأشارت النتائج إلى وجود فروق دالة إحصائياً عند مستوى دلالة (0,01) بين المعلمات في مدى المعرفة لصالح مجموعة البكالوريوس، بينما أشارت النتائج إلى عدم وجود فروق دالة في مدى المعرفة تعزى للتخصص أو لسنوات الخبرة التعليمية، كما أشارت النتائج إلى وجود فروق دالة إحصائياً عند مستوى دلالة (0.01) في ممارسة الكفاءات الخاصة بإعداد وتصحيح وتحليل الاختبارات لصالح حملة البكالوريوس، كما وجدت فروق دالة إحصائياً عند مستوى دلالة (0,05) في ممارسة المعلمات لصياغة الأهداف السلوكية وتصحيح الاختبارات لصالح فئة الخبرة (11سنة فأكثر)، ووجدت فروق دالة إحصائياً عند مستوى دلالة (0,05) في ممارسة المعلمات في سرعة إعادة الأوراق بعد التصحيح، واستخدام عبارات التشجيع لصالح فئة الخبرة الأكثر، كما وجدت فروق دالة إحصائياً عند مستوى دلالة (0,05) في استخدام اختبار المقال لصالح فئة الخبرة(6-10)، وأخيراً أشارت النتائج إلى عدم وجود فروق دالة إحصائياً في مدى ممارسة المعلمات في إعداد وتصحيح الاختبارات تعزى للتخصص (جغرافيا و تاريخ).

وهدفت دراسة الغمري (1998) إلى التعرف على مستوى معرفة معلمي المرحلة الأساسية في اليمن بمهارات تطوير الاختبارات المدرسية، واتجاهاتهم نحوها، وتكونت عينة الدراسة من(525) معلماً من معلمي المرحلة الأساسية في أمانة العاصمة، وقد استخدمت أداتين الأولى: هي اختبار لقياس كفايات المعلمين بمهارات تطوير الاختبار الذي أعده جرادات(1988)، والثانية: هي مقياس اتجاه المعلمين نحو الامتحانات المدرسية الذي طوره عودة (1990)، وأشارت النتائج إلى أن مستوى المعرفة لدى معلمي المرحلة الأساسية لم تصل إلى المستوى المقبول تربوياً (80%)، كما أشارت النتائج إلى ان معرفة المعلمين ذوي الخبرة القصيرة (5 سنوات فأقل) أعلى من معرفة المعلمين ذوي الخبرة الطويلة (أكثر من 5 سنوات)، كما أن معرفة المعلمين من حملة البكالوريوس كانت أعلى من معرفة المعلمين من

حملة المؤهلات الثلاثة الأخرى (دبلوم متوسط، معهد معلمين، ثانوية عامة)، كما أشارت النتائج من جهة ثانية إلى أن المعلمين أبدوا اتجاهات إيجابية نحو الاختبارات المدرسية، وأن الجنس والخبرة ليس لهما أثر على الاتجاهات، أمّا بالنسبة للمؤهل فأشارت النتائج إلى أن المعلمين من حملة البكالوريوس لديهم اتجاهات أكثر إيجابية من حملة المؤهلات الثلاث الأخرى.

خلاصة الدراسات السابقة:

ومن خلال استعراضنا للدراسات السابقة يتضح أن الدراسة الحالية قد اتفقت مع بعض من الدراسات السابقة من حيث سعيها إلى التعرف على مدى معرفة، أو فهم المعلمين للمهارات اللازمة لتطوير الاختبارات المدرسية مثل دراسة كيربر (Kerber,1976) ودراسة نيومن (Newman,1981) ودراسة جرادات (1988) ودراسة بيج ومارسو (Marso&Pigge,1988)، كما وأن الدراسة الحالية قد تشابهت مع بعض الدراسات السابقة في نوعية الأداة المستخدمة في جمع البيانات مثل دراسة كيربر (Kerber,1976)ودراسة جرادات (1988) حيث تم استخدام اختبار لقياس مهارات المعلمين في تطوير الاختبارات المدرسية، إذ أن باقي الدراسات استخدمت استبانة أو قائمة من المعايير.

هذا وقد دلت نتائج الدراسات التي بحثت في مدى معرفة المعلمين بمهارات تطوير الاختبارات أن هذه المعرفة تتراوح بين الضعيف والمتوسط، ومن هنا جاءت هذه الدراسة لتحديد مدى معرفة المعلم الأردني في مجال بناء الاختبارات التحصيلية، وهل يختلف المعلم الأردني عن غيره من المعلمين في الخارج؟ ولبيان هذا فقد تم تطوير اختبار لقياس معارف المعلمين في مجال بناء اختبارات التحصيل.

الفصل الثالث

المنهجية والتصميم

يتضمن هذا الفصل وصفاً لمجتمع الدراسة وعينتها وطريقة اختيارها، والأداة التي جرى تطويرها لأغراض التعرف على كفايات المعلمين في مجال بناء اختبارات التحصيل، كما ويتضمن وصفاً لإجراءات الدراسة التي تتألف من خطوات تطوير الاختبار وتطبيقه وتصحيحه، واشتقاق مؤشرات ثباته وصدقه ومعايير تفسير الأداء على فقراته.

1.3 مجتمع الدراسة:

يتكون مجتمع الدراسة من كافة معلمي المرحلة الثانوية في مديريات التربية والتعليم التابعة لوزارة التربية والتعليم في الأردن، البالغ عددها(33) مديرية موزعة على ثلاثة أقاليم هي(الشمال، الوسط، الجنوب). والجدول رقم(1) يوضح أعداد المعلمين حسب الجنس والأقاليم ومديريات التربية، كما جاءت في إحصائيات وزارة التربية والتعليم للفصل الدراسي الأول لعام 2005/2004 .

جدول رقم(1)

توزيع أفراد مجتمع الدراسة حسب الإقليم ومديرية التربية

الإقليم	مديرية التربية	العدد		المجموع
		ذكور	إناث	
	اربد الأولى	518	490	
	اربد الثانية	132	179	
	اربد الثالثة	58	86	
	لواء الكورة	155	213	
	لواء الأغوار ش	122	105	
	لواء بني كنانة	168	122	
	قصبة المفرق	160	182	
	لواء الباديةش ش	110	84	4025
	لواء الباديةش غ	117	143	
	محافظة جرش	196	225	
الشمال	محافظة عجلون	208	252	
	عمان الأولى	604	420	
	عمان الثانية	367	345	
	عمان الثالثة	293	302	
	عمان الرابعة	374	428	
	البادية الوسطى	82	79	
	قصبة ماديا	150	168	
	لواء ذيبان	41	50	
	قصبة الزرقاء	369	437	
	لواء الرصيفة	189	241	5905
	قصبة السلط	204	297	
	لواء دير علا	93	64	
	لواء الشونه ج	66	62	
الوسط	عين الباشا	78	102	
	قصبة الكرك	180	212	
	لواء المزار ج	238	124	
	لواء القصر	71	57	
	لواء الأغوار ج	48	53	
	محافظة الطفيلة	181	161	
	قصبة معان	127	314	
	لواء البتراء	47	70	2119
	لواء الشوبك	10	11	
الجنوب	محافظة العقبة	92	123	
المجموع		5848	6201	12049

2.3 عينة الدراسة:

ولتحديد أفراد العينة تم تقسيم المملكة إلى ثلاثة أقاليم، وذلك حسب التقسيم الإداري المعتمد وهي على النحو التالي:

1. إقليم الشمال ويتكون من إحدى عشرة مديرية وهي: اربد الأولى، اربد الثانية،اربد الثالثة، بني كنانة، الكورة، الأغوار الشمالية، جرش،عجلون،المفرق،البادية الشمالية الشرقية، البادية الشمالية الغربية.
2. إقليم الوسط ويتكون من ثلاث عشرة مديرية وهي :عمان الأولى، عمان الثانية،عمان الثالثة،عمان الرابعة، السلط، ديرعلاّ، الشونه الجنوبية عين الباشا، مأدبا، ذيبان، الزرقاء، الرصيفة.
3. إقليم الجنوب ويتكون من تسع مديريات وهي:قصبة الكرك، المزار الجنوبي، القصر، الأغوار الجنوبية، الطفيلة، معان، الشوبك، البتراء،العقبة.

وقد تم اختيار عينة عشوائية عنقودية وعلى مراحل، حيث تم الاختيار العشوائي لمديرية واحدة من كل إقليم، ثم تم الاختيار العشوائي لـــ(5)مدارس ثانوية من الذكور و (5)مدارس ثانوية من الإناث من نفس المديرية، فيصبح عدد المدارس المختارة هي (15) للذكور، و(15) للإناث، وتم تطبيق الاختبار على المعلمين والمعلمات ضمن هذه المدارس. والجدول رقم (2) يوضح عدد أفراد العينة التي تم اختيارها من الأقاليم الثلاثة حسب المديرية المختارة.

جدول رقم (2)

توزيع أفراد العينة حسب الإقليم ومديرية التربية

الإقليم	مديرية التربية	العدد		المجموع
		ذكور	إناث	
الشمال	محافظة جرش	103	96	199
الوسط	قصبة الزرقا	121	112	233
الجنوب	قصبة الكرك	98	107	205
	المجموع	322	315	637

3.3 أداة الدراسة:

تم بناء اختبار لقياس مدى تمكن المعلمين من كفايات بناء الاختبار التحصيلي ، وبما أن الهدف من بناء هذا الاختبار ليس المقارنة بين المعلمين، إذ لا يوجد محتوى محدد تعرض له المعلمون يراد اختبارهم فيه، ولذلك فإن بناء اختبار محكي المرجع قد بدا مناسباً لهذا الغرض.

ويتطلب بناء اختبار من هذا النوع عدة إجراءات وهي:

1.3.3 تحديد الغرض من الاختبار:

إن الهدف من الاختبار هو قياس كفايات المعلمين في مجال بناء اختبارات التحصيل، لقياس مدى تمكنهم من المهارات الأساسية والفرعية التي سيصار إلى تحديدها، والكشف عن مدى اكتسابهم لهذه المهارات، ويمكن عن طريقه تشخيص جوانب القوة والضعف في هذا المجال، مما قد يساعد صانعي القرار في المؤسسات التربوية في اتخاذ القرارات المناسبة.

2.3.3 تحديد المجال أو(الإطار):

تم تحديد المجالات التي يمكن أن تغطي ماُ يحتاجه المعلم عند بنائه لاختبارات التحصيل الصفية التي يضعها بنفسه اعتماداً على المنحى النظري كطريقة لاشتقاق المجالات الخاصة ببناء الاختبارات، وبناء على افتراضات هذا المنحى تمت مراجعة كثير مما كتب عن الاختبارات من حيث الإعداد وكتابة الفقرات والإخراج والتطبيق والتصحيح وتفسير النتائج وصياغة الأهداف.

ومن الكتب والمراجع والدراسات التي اعتمد عليها في إعداد المجالات ما يلي:(عودة 1990؛عودة 1993؛ لندفل1968؛عدس1988؛جامل1995؛علام1985)
(Ebel 1991؛Gronlund 1976؛Gay 1980؛Hopkins and Antes 1978)

ولقد تم التوصل إلى ستة مجالات يمكن ان تغطي ما يحتاجه المعلم في مجال بناء اختبارات التحصيل، وهي على النحو التالي:

1) الأهداف التعليمية وصياغتها ودورها في عملية بناء اختبارات التحصيل.
2) جدول المواصفات ودوره في بناء اختبارات التحصيل.

3) أنواع الفقرات الاختبارية وصياغتها.

4) خصائص الفقرة الاختبارية الجيدة وانتقاؤها وتهيئة الاختبار بشكله النهائي.

5) خصائص الاختبار الجيد.

6) تصحيح الاختبار وتفسير نتائجه.

3.3.3 أهداف المجالات (الأهداف المرحلية):

حدد لكل مجال من المجالات التي سبق ذكرها المعلومات اللازمة للمعلم في عملية بنائه لاختبارات التحصيل، ثم قدمت المجالات الستة، والأهداف النهائية التي تغطي كل مجال إلى مجموعة من المحكمين من اختصاصي القياس والتقويم في كل من جامعة مؤتة، والجامعة الأردنية بغرض معرفة تمثيل المجالات الستة لمحتوى الاختبار الذي يهدف إلى قياسه، وانتماء الأهداف المرحلية (النهائية) لكل مجال من هذه المجالات، ومدى صلاحيتها لتغطية المجالات وفقاً لحاجة المعلمين ذات العلاقة بهذه المجالات، واقتراح التعديلات التي يرونها مناسبة، ونتيجة لذلك أصبحت الأهداف النهائية كما هي موضحة بالجدول رقم (3).

جدول رقم (3)

تحليل المجالات إلى أهداف مرحلية

المجال	الأهداف المرحلية
1) معرفة المعلم للأهداف التعليمية ودورها في عملية التعلم والتعليم وفقا لأصنافها.	• أن يميز المعلم بين الهدف التربوي والتعليمي والسلوكي • أن يميز المعلم بين الأهداف المعرفية والانفعالية والمهارية • أن يحدد أوجه النقص في صياغة الهدف السلوكي، وفقاً لشروط صياغته. • أن يصنف أهدافا معرفية سلوكية وفقاً لتصنيف بلوم.
2) معرفة المعلم بجدول المواصفات ودوره في بناء اختبارات التحصيل	• أن يعرف مكونات جدول المواصفات. • أن يعرف إجراءات إعداد جدول المواصفات. • أن يعرف كيفية اشتقاق الأوزان النسبية لمكونات جدول المواصفات (المحتوى،الأهداف، المستويات، عدد الفقرات) • أن يعرف كيفية تحديد إعداد الفقرات في كل خلية .
3) معرفة المعلم بكيفية اختيار الفقرات المناسبة لقياس أهدافه	• أن يحدد مزايا كل نوع من أنواع الفقرات الإختبارية. • أن يحدد عيوب كل نوع من أنواع الفقرات الإختبارية. • أن يعين الأهداف المعرفية أو مستويات التفكير التي يصلح لها كل نوع من أنواع الفقرات، وينتقي الفقرات المناسبة للهدف. • أن يعرف خصائص الفقرة الجيدة لكل نوع من أنواع الفقرات الإختبارية. • أن يقترح التعديل المناسب إذا أعطيت له فقرة غير جيدة.
4) معرفة المعلم بطرق انتقاء فقرات الاختبار وتهيئة بشكله النهائي	• أن يطبق طرق تحليل الفقرات (معامل الصعوبة، معامل التمييز، فعالية البدائل) • أن يعرف كيفية ترتيب الفقرات وفقاً لتحليل الفقرة أو محتواها أو نوعها. • أن يعرف شروط كتابة تعليمات الاختبارات التحصيلية.
5) معرفة المعلم بخصائص الاختبار الجيد	• أن يميز بين مفهومي الصدق والتمييز. • أن يحدد الإجراءات التي يمكن بواسطتها الحصول على مؤشرات عن الصدق. • أن يحدد الإجراءات التي يمكن بواسطتها الحصول على مؤشرات عن الثبات في ضوء أغراض الاختبار.
6) معرفة المعلم بتصحيح الاختبار وتفسير نتائجه.	• أن يعرف طرق تصحيح اختبارات التحصيل بنوعيها، المقالية والموضوعية. • أن يتعرف على الفرق بين الطريقة الكلية و التحليلية في تصحيح الأسئلة المقالية. • أن يقدم وصفاً للبيانات من خلال مقاييس النزعة المركزية والتشتت والمئينات والدرجة المعيارية. • أن يفسر الدرجة الخام بالطرق الإحصائية. • أن يميز بين التفسير المحكي المرجع والتفسير المعياري المرجع.

4.3.3 تحليل الأهداف المرحلية النهائية إلى أهداف تفصيلية سلوكية:

في هذه الخطوة حُللت الأهداف النهائية لكل مجال من مجالات الاختبار الرئيسية إلى أهداف أكثر تفصيلاً، لأن الأولى تعبر عن نواتج تعليمية مركبة قد يصعب قياسها مباشرة، وذلك باستخدام طريقة جانيه(Gagne) في تحليل المهام (Task Analysis)، بحيث يمكن قياس كل هدف بفقرة اختبارية أو أكثر، ونظمت على هيئة مستطيلات تمثل أهدافا أساسية، وعلى هيئة دوائر تمثل أهدافا مساعدة لتحقيق الأهداف الأساسية، إذ لا يستطيع المتعلم اكتساب الهدف الأساسي، إلا بعد مروره أو معرفته بالهدف المساعد الممهد له، وتفيد هذه الأطر التخطيطية في تحديد المعلومات اللازمة التي تهدف إلى تحقيق الأهداف النهائية للاختبار، كما هي موضحة بالمخططات [انظر الملحق (أ)].

5.3.3 بناء مفردات الاختبار:

على ضوء الأهداف التفصيلية التي أعدت في المرحلة السابقة، تم الاطلاع على بعض الاختبارات التي تم إعدادها وتطويرها في هذا المجال، وبعض نماذج الاختبارات المنشورة في كتب القياس والتقويم، فقد تمت صياغة (140) فقرة من نوع الاختيار من متعدد بأربعة بدائل، تغطي الأهداف النهائية، وأهدافها التفصيلية للمجالات الستة للاختبار، وبواقع فقرة على الأقل لكل هدف تفصيلي، وبعد مراجعتها اختيرت من بينها الفقرات ذات الصياغة الجيدة والتي تقيس أهدافا تفصيلية أساسية أو ثانوية.

وقد اتبعت هذه الخطوة نظراً لان إعطاء اختبار طويل (140)فقرة يكون متعبا ويقود إلى الملل، ويقلل من دافعية المستجيب لإكمال الاختبار، وبالتالي يفقد الاختبار أهم خاصية، وهي الاختبار تحت ظروف أقصى الأداء، وبذلك اصبح الاختبار في صورته الأولية يشتمل على(80)فقرة، اي بنسبة 57% من مجموع الفقرات التي غطت مجالات الاختبار الستة والموضحة في الأطر التخطيطية.

وقد وزعت فقرات الاختبار على المجالات الستة، والتي يوضحها الجدول التالي:

جدول(4)

توزيع أعداد الفقرات على مجالات الاختبار

المجالات	عدد الفقرات	النسبة المئوية
الأهداف التعليمية وصياغتها ودورها في بناء اختبارات التحصيل	16	%20
جدول المواصفات ودوره في بناء اختبارات التحصيل	6	%7.5
أنواع الفقرات الاختبارية وصياغتها	23	%28.75
انتقاء فقرات الاختبار وتهيئته بشكله النهائي	11	%13.75
خصائص الاختبار الجيد	10	%12.5
تصحيح الاختبار وتفسير نتائجه	14	%17.5
مجموع الفقرات	80	%100

هذا وقد حددت المستويات المعرفية التي يقيسها الاختبار بثلاثة مستويات (معرفة،استيعاب، تطبيق فما فوق)، استناداً إلى تصنيف بلوم وإمكانية الدمج بين مستوياته.

وقد تم توزيع فقرات الاختبار وفقاً للمستويات المعرفية التي حددت سابقاً وبالنسب (54%، 30 %، 16%)على الترتيب ويلاحظ أن ما خصص لقياس المعرفة حصل على أعلى النسب كون الاختبار يهدف إلى قياس معارف المعلمين في مجال بناء اختبارات التحصيل، وتلى ذلك الاستيعاب ثم التطبيق.

6.3.3 تجهيز الاختبار بصيغته الأولية وتجريبه وتصحيحه:

رتبت فقرات الاختبار في صورته الأولية وفقاً لتتابع مجالاته الستة، وفي كل مجال رتبت الفقرات بحسب التقدير الأولي للصعوبة، وزود الاختبار بورقة إجابة موضحاً فيها هدف الاختبار، وطريقة الإجابة بوضع علامة(×) أمام البديل الصحيح في ورقة الإجابة. ثم طبق الاختبار على (30)طالب وطالبة من طلبة الماجستير في القياس والتقويم وجميعهم درس مقرر في القياس والتقويم، والمتضمن بناء اختبارات التحصيل وعلى مدى فصل دراسي كامل، وعدت هذه المجموعة هي المجموعة المدربة، ولقد تم الإشراف على تطبيق الاختبار وتهيئة الجو المناسب والموحد

للطلبة، وتحت ظروف أقصى أداء. كما وطبق الاختبار على(30) معلماً من المعلمين العاملين في وزارة التربية والتعليم ويمارسون وضع الاختبارات التحصيلية وموزعين على (3) مراكز كانت مخصصة للتدريب على دورة في الحاسوب، هذا وقد عدت مجموعة المعلمين التي لم تتلقى التدريب فيما يتعلق بمجالات الاختبار هي المجموعة غير المدربة، وبعد الانتهاء من إجراءات التطبيق تم تصحيح أوراق الإجابة حيث استخدم مفتاح تصحيح مثقب خاص بالاختبار، وتم اتباع التعليمات التالية عند التصحيح:

1) فحص العلامات (الإشارات) التي وضعها المفحوصون في الأوراق الخاصة بالإجابة للتأكد من وجود علامة واحدة أمام كل فقرة، وفي حال وجود اكثر من علامة واحدة أمام الفقرة تعتبر لاغية حتى إذا كانت إحدى العلامات هي الإجابة الصحيحة.
2) يطبق مفتاح التصحيح –بعد ثقب الأماكن المشار إليها بعلامة(×)، والتي تطابق تماماً المربعات على ورقة الإجابة –على ورقة الإجابة ويضبط المفتاح على هذه الورقة.
3) إعطاء درجة للإجابة الصحيحة عن كل فقرة، وصفر للإجابة الخاطئة أو عند ترك المفحوص الفقرة دون إجابة. وبالتالي فإن درجة المفحوص الكلية على الاختبار تساوي مجموع درجاته على الفقرات الثمانين(80)، وتتراوح الدرجة بين (0 – 80).

7.3.3 تحليل فقرات الاختبار:

الطريقة التي اتبعت في تحليل فقرات الاختبار محكي المرجع، هي الطريقة التي تعتمد على قياس مجموعتين مختلفتين من الأفراد، إحداهما تلقت التدريب فيما له علاقة بمجالات الاختبار الستة، والأخرى لم تتلق التدريب، حيث تعتمد هذه الطريقة في حساب(معامل حساسية الفقرة) على طرح عدد الأفراد الذين أجابوا على المفردة إجابة صحيحة في المجموعة التي تلقت التعليم، من عدد الأفراد الذين أجابوا عليها إجابة صحيحة في المجموعة التي لم تتلقى التعليم مقسوماً على عدد

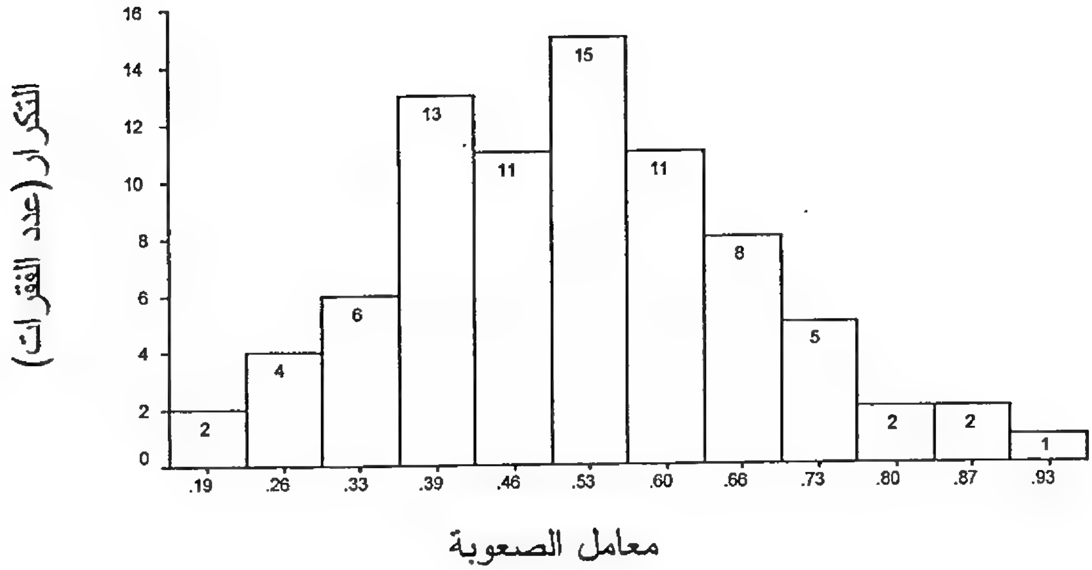
أفراد إحدى المجموعتين، أما معاملات الصعوبة فتم تحديدها بإيجاد النسبة لعدد الأفراد الذين أجابوا على المفردة إجابة صحيحة في كلا المجموعتين من مجموع الذين سوف يجرب عليهم الاختبار.

وعلى ضوء ما سبق حسبت معاملات الصعوبة والتمييز لفقرات الاختبار (80)في صورته الأولية، وذلك بتطبيق المعادلات الخاصة بإيجاد معاملات التمييز، ومعاملات الصعوبة على فقرات الاختبار من مجموعة عينة التجريب المؤلفة من 30 ورقة إجابة للمجموعة المدربة و30 ورقة إجابة للمجموعة المدربة وبذلك حسبت معاملات الصعوبة والتمييز لفقرات الاختبار لـــ (60)ورقة كعينة لتجريب الاختبار، وكما يوضحها الجدول رقم (5) والأشكال (1)، (2)

جدول (5)

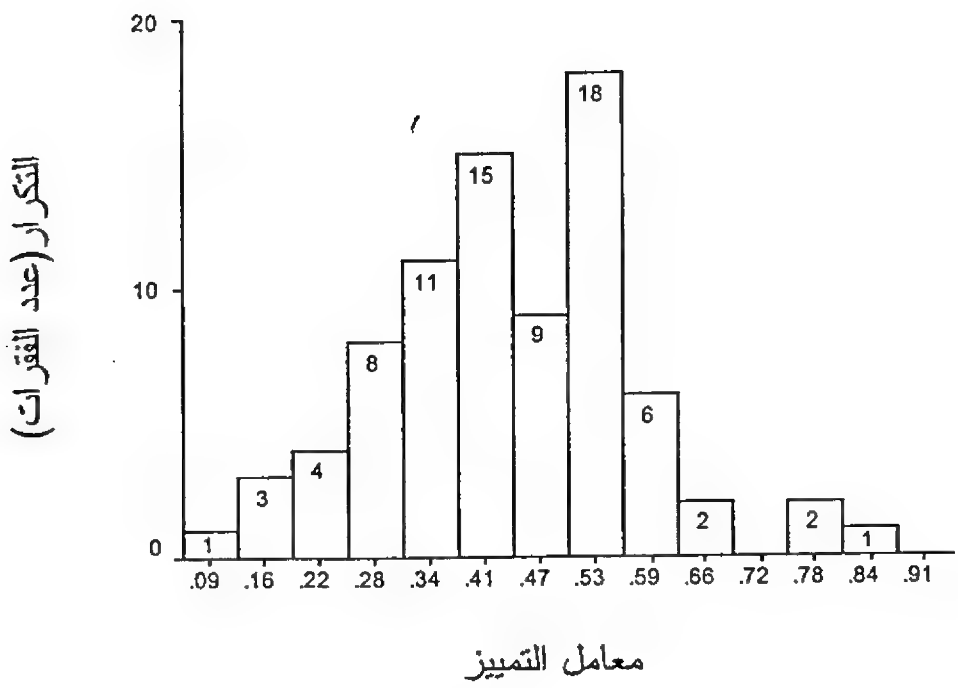
معاملات الصعوبة والتمييز لفقرات الاختبار في صورته الأولية

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
1	0.7	0.5	28	0.6	0.5	55	0.22	0.35
2	0.93	0.15	29	0.5	0.4	56	0.2	0.15
3	0.67	0.45	30	0.6	0.5	57	0.5	0.4
4	0.83	0.35	31	0.77	0.35	58	0.55	0.4
5	0.53	0.65	32	0.72	0.55	59	0.45	0.80
6	0.45	0.5	33	0.73	0.55	60	0.67	0.55
7	0.53	0.35	34	0.68	0.35	61	0.7	0.6
8	0.37	0.2	35	0.25	0.2	62	0.37	0.25
9	0.6	0.6	36	0.4	0.4	63	0.38	0.45
10	0.48	0.15	37	0.5	0.5	64	0.6	0.4
11	0.48	0.45	38	0.5	0.4	65	0.5	0.6
12	0.42	0.55	39	0.62	0.55	66	0.72	0.4
13	0.27	0.35	40	0.33	0.35	67	0.65	0.4
14	0.35	0.5	41	0.43	0.45	68	0.58	0.75
15	0.6	0.6	42	0.6	0.4	69	0.43	0.45
16	0.33	0.45	43	0.45	0.4	70	0.55	0.6
17	0.3	0.35	44	0.5	0.4	71	0.53	0.4
18	0.43	0.55	45	0.65	0.4	72	0.42	0.55
19	0.65	0.3	46	0.67	0.65	73	0.35	0.6
20	0.5	0.2	47	0.53	0.55	74	0.42	0.45
21	0.53	0.55	48	0.42	0.35	75	0.47	0.85
22	0.6	0.5	49	0.4	0.5	76	0.85	0.3
23	0.85	0.3	50	0.25	0.2	77	0.38	0.45
24	0.63	0.35	51	0.23	0.30	78	0.6	0.30
25	0.65	0.4	52	0.32	0.35	79	0.38	0.45
26	0.4	0.3	53	0.45	0.4	80	0.42	0.55
27	0.52	0.05	54	0.43	0.25			



الشكل رقم (1)

توزيع فقرات الاختبار حسب مستوى الصعوبة في العينة التجريبية

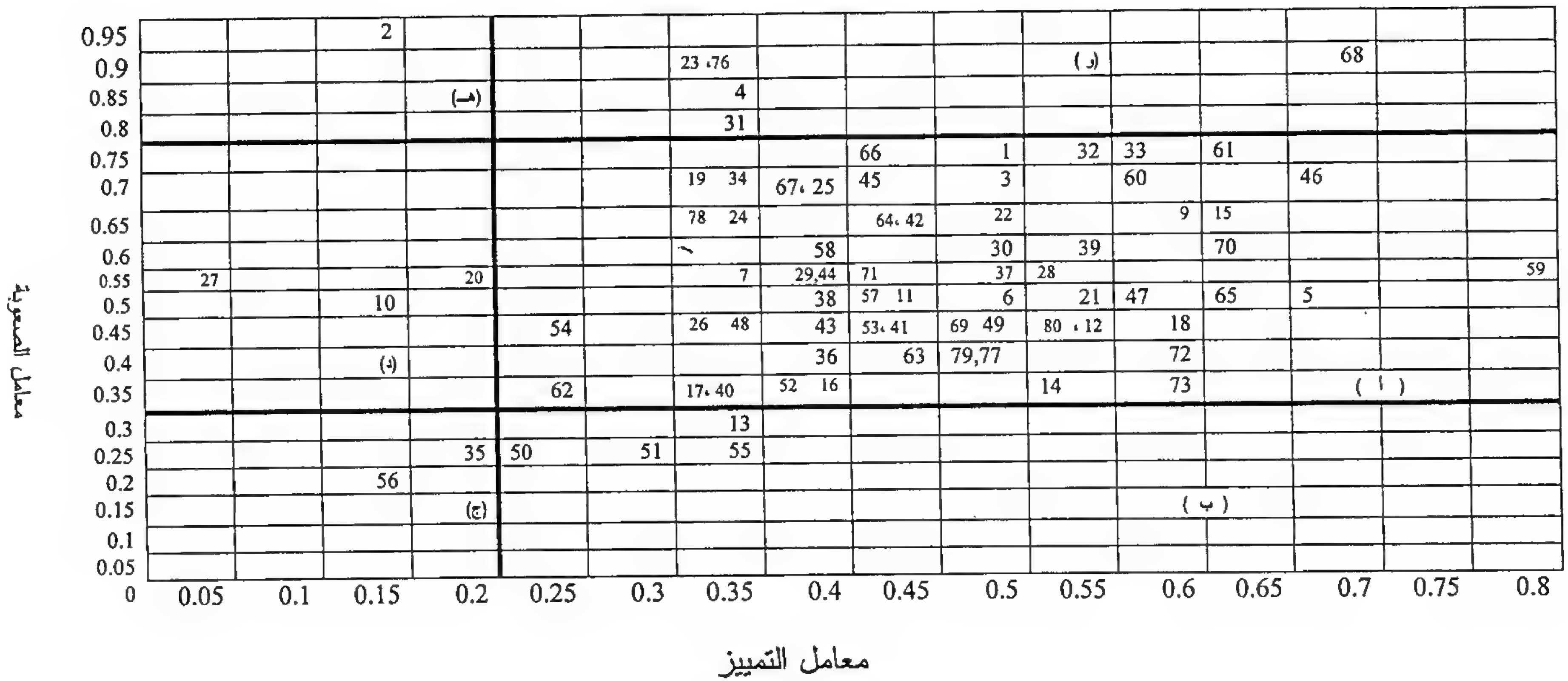


الشكل رقم (2)

توزيع فقرات الاختبار حسب معامل التمييز

8.3.3 انتقاء فقرات الاختبار بصورته النهائية :

على ضوء نتائج عملية تحليل الفقرات تم تمثيل (80) فقرة في مخطط انتشار مكوّن من محورين : أفقي يمثل معاملات تمييز الفقرات، وعامودي يمثل معاملات الصعوبة، والشكل (3) يوضح ذلك.



شكل (3)

مخطط انتشار الفقرات وفقا لمعاملات الصعوبة والتمييز

من الجدول (5) والشكل (3) يتبين ما يلي:

1) قبول جميع الفقرات التي تقع في المساحة (أ) كونها ذات معاملات صعوبة وتمييز مقبولين. وقد استخدم المعيار (0.2)حسب ما أوضحه إيبل و فرسبي (Ebel&Frisbe,1991)، الذي يصف الفقرات الحاصلة على تمييز (0.2)فأكثر على أنها فقرات جيدة. واستخدم المعيار (0.3 – 0.75) لاختيار الفقرات وفقا لمعامل الصعوبة.

2) استبعدت جميع الفقرات التي تقع في المساحات المتبقية (ب)، (ج)، (د)، (هـ) لأنها إمّا ينقصها معامل التمييز الجيد أو معامل الصعوبة الجيد، وهي الفقرات ذات الأرقام التالية(2، 4، 8، 10، 13، 20، 23، 27، 31، 35) عدا بعض الفقرات التي تغطي محتوى لم تغطيه أي من فقرات المساحة (أ) وهذا ما يعزز صدق المحتوى وهي الفقرات (50، 51 ، 54، 56,55، 62).

وبناءً على ما سبق، فقد اشتمل الاختبار في صورته النهائية على (70) فقرة اختبارية وكما في الجدول (6) الذي يوضح معاملات صعوبتها وتمييزها.

جدول (6)

معاملات الصعوبة والتمييز للفقرات النهائية للاختبار

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
1	0.7	0.5	25	0.68	0.35	49	0.45	0.80
2	0.67	0.45	26	0.4	0.4	50	0.67	0.55
3	0.53	0.65	27	0.5	0.5	51	0.7	0.6
4	0.45	0.5	28	0.5	0.4	52	0.37	0.25
5	0.53	0.35	29	0.62	0.55	53	0.38	0.45
6	0.6	0.6	30	0.33	0.35	54	0.6	0.4
7	0.48	0.45	31	0.43	0.45	55	0.5	0.6
8	0.42	0.55	32	0.6	0.4	56	0.72	0.4
9	0.35	0.5	33	0.45	0.4	57	0.65	0.4
10	0.6	0.6	34	0.5	0.4	58	0.58	0.75
11	0.33	0.45	35	0.65	0.4	59	0.43	0.45
12	0.3	0.35	36	0.67	0.65	60	0.55	0.6
13	0.43	0.55	37	0.53	0.55	61	0.53	0.4
14	0.65	0.3	38	0.42	0.35	62	0.42	0.55
15	0.53	0.55	39	0.4	0.5	63	0.35	0.6
16	0.6	0.5	40	0.25	0.2	64	0.42	0.45
17	0.63	0.35	41	0.23	0.30	65	0.47	0.85
18	0.65	0.4	42	0.32	0.35	66	0.85	0.3
19	0.4	0.3	43	0.45	0.4	67	0.38	0.45
20	0.6	0.5	44	0.43	0.25	68	0.6	0.30
21	0.5	0.4	45	0.22	0.35	69	0.38	0.45
22	0.6	0.5	46	0.2	0.15	70	0.42	0.55
23	0.72	0.55	47	0.5	0.4			
24	0.73	0.55	48	0.55	0.4			

9.3.3 مؤشرات الصدق:

يعتبر الصدق من أهم الخصائص السيكومترية في الاختبارات، ويعني الصدق أن الاختبار قد صمم لقياس سمة خاصة أو سلوك معين، وأن فقرات هذا الاختبار ترتبط كلها بما يراد قياسه، أي أن الاختبار الصادق هو الذي يقيس السمة أو السلوك الذي وضع الاختبار من أجل قياسه، وبذلك يعطي هذا الاختبار درجة دقيقة تمثل موقع المفحوص على خط الأعداد الخاص بهذا السلوك.

وبما أن هنالك عدة معاني أو مصطلحات تتعلق بصدق الاختبار، فإن هنالك أكثر من مؤشر أو طريقة للكشف عن صدق الاختبار، وفي هذه الدراسة تم إتباع مؤشرين لتقدير صدق الاختبار وهما على النحو التالي:

1) أولاً: صدق المحتوى:

يرتبط تحقيق صدق المحتوى (المضمون) بطريقة وضع الاختبار، ويعتبر صدق المحتوى أهم من غيره من أنواع الصدق الأخرى عندما يكون المجال السلوكي الذي يقيسه الاختبار محدداً. ويشير علاّم(1995)إلى أن إتباع الخطوات المناسبة في بناء الاختبارات محكية المرجع سيؤدي بالضرورة إلى اختبارات صادقة دون الحاجة إلى مزيد من التحقق.

2) ثانياً: الصدق التمييزي:

تم اعتماد طريقة المقارنة بين أداء مجموعتين من الأفراد على الاختبار، والتي أشار إليها علاّم (1995) وتهدف إلى المقارنة بين أداء مجموعتين من الأفراد على الاختبار الذي يقيس المعارف والمعلومات التي يتضمنها كل مجال من مجالات الاختبار الستة، ونختار مجموعتين من الأفراد بحيث تكون إحداهما تلقت التعليم أو التدريب فيما يتعلق بمجالات الاختبار، والثانية لم تتلق التعليم.

في ضوء ذلك فقد قدم الاختبار في صورته الأولية إلى مجموعتين، إحداهما تلقت التعليم (طلبة الماجستير في القياس والتقويم والذين تلقوا تعليماً منتظما في بناء اختبارات التحصيل)، والمجموعة الثانية لم تتلق تعليماً منتظماً شاملاً في مجالات الاختبار الستة، وهم (مجموعة من المعلمين العاملين في وزارة التربية والتعليم ويمارسون وضع الاختبارات التحصيلية للعام الدراسي (2006/2005).

10.3.3 تقدير ثبات الاختبار:

يشير مفهوم ثبات الاختبار إلى دقته في قياس السمة التي صمم لقياسها، ويعد ثبات الاختبار من المقومات الأساسية للاختبار الجيد، حيث يفترض أن يعطي الاختبار نفس النتائج تقريباً، أي استقرار النتائج فيما إذا أعيد استخدامه في أوقات مختلفة وظروف متشابهة، ويعرّف الثبات إحصائياً، بأنه نسبة التباين الحقيقي إلى التباين الكلي، أو مربع معامل الارتباط بين العلامات الحقيقية والظاهرية (عودة، 1993).

وقد تم استخدام مؤشرين لتقدير ثبات الاختبار:

1) حساب معامل الثبات الطريقة النصفية (Split-half Reliability)

حيث تقوم فكرة معامل الثبات بهذه الطريقة على حساب معامل الارتباط بين علامات معلمي عينة الدراسة على فقرات نصف الاختبار، وعلاماتهم على فقرات النصف الآخر من الاختبار.

2) حساب معامل الثبات (Kr-20):

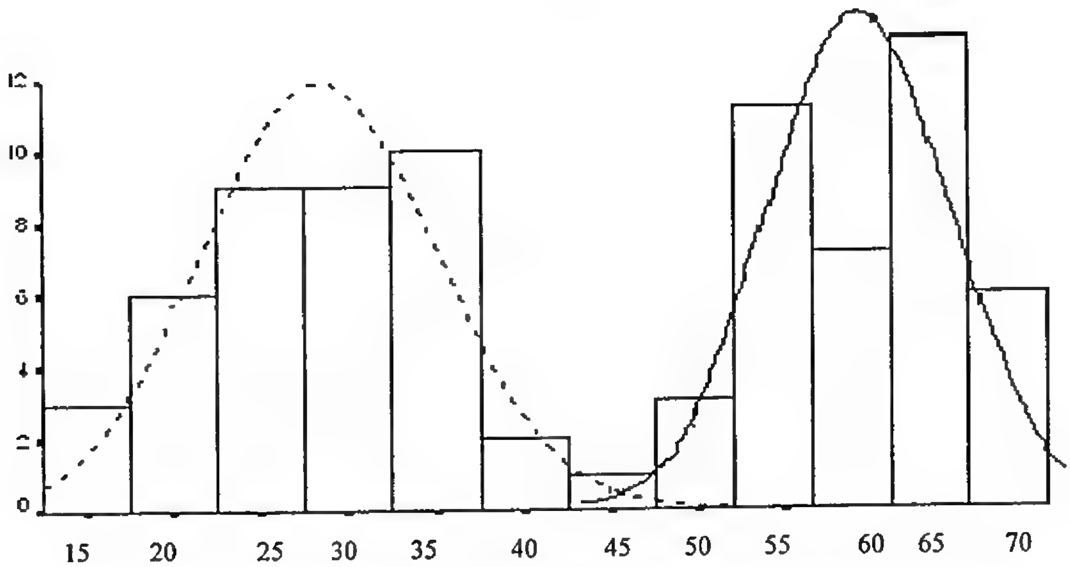
في هذه الطريقة تم حساب معامل الارتباط بين العلامات لمجموعة الثبات على جميع الفقرات الداخلة في الاختبار، وكأننا بذلك لانقسم الاختبار إلى نصفين كما هو متبع بالطريقة النصفية، بل هو تقدير لجميع معاملات الأنصاف المحتملة، وهي هنا (70) فقرة.

11.3.3 تقدير مستوى الإتقان (المحك) على الاختبار :

لتقدير مستوى الأداء على الاختبار، استخدمت طريقة المجموعات المتناقضة (Contrasting groups method)، وتقوم هذه الطريقة على اختيار مجموعتين من الأفراد إحداهما مدربة والأخرى غير مدربة على الكفاية المطلوبة، ثم يطبق الاختبار على كلا المجموعتين، ويُرسم منحنيان يمثلان عدد طلاب المجموعتين الذين حصلوا على درجات كلية مختلفة في الاختبار، ويمكن اعتبار نقطة تقاطع المنحنيين بمثابة قيمة تقديرية للمحك الذي يمثل مستوى الأداء المطلوب(علام، 1995).

وقد اتبعت هذه الطريقة دون غيرها، كونها أكثر مناسبة في ظروف هذا الاختبار، وينصح بيرك(Berk) بالاعتماد على طريقة المجموعات المتناقضة إذ إنها تعد أكثر الطرق ملائمة وواقعية، ويسراً للمحكمين في انتقاء الأفراد عند طرفي متصل الكفاية (علاّم ،1995). وقد عدت مجموعة طلبة الماجستير الذين طبق عليهم الاختبار، والذين تلقوا تعليماً منتظماً في مجالات الاختبار هم (المجموعة المدربة)، في حين عدت مجموعة المعلمين والذين لم يتلقوا تعليماً منتظماً فيما يتعلق بمجالات الاختبار، هم المجموعة غير المدربة، وقد طبق الاختبار على كلا

المجموعتين بشكل منفصل، ويوضح الشكل رقم (4) التمثيل البياني للدرجات النهائية التي أحرزتها المجموعتين (40 فرداً للمجموعة المدربة و40 فرداً للمجموعة غير المدربة) على الاختبار في صورته النهائية.



شكل رقم (4)

التمثيل البياني للدرجات النهائية للمجموعتين بغرض تقدير مستوى الإتقان على الاختبار (درجة التمكن)

ومن الشكل (4) نجد أن نقطة تقاطع المنحنيين للمجموعتين (المدربة وغير المدربة) هي الدرجة (45)، وبذلك تكون هي الدرجة التي تحدد مستوى الإتقان على هذا الاختبار (الاختبار المعرفي لقياس كفايات المعلمين في مجال بناء اختبارات التحصيل)، وبالتالي فإن أي فرد في المجموعتين التي طبق عليهما الاختبار يعد متمكناً من المعارف التي يقيسها الاختبار إذا حصل على درجة (45) وهي تمثل نسبة (64%) من أهدافه.

12.3.3 : تطبيق الاختبار :

بعد الانتهاء من إجراءات بناء الاختبار، وتحديد مستوى الإتقان على الاختبار، تم تطبيق الاحتبار على افراد العينة وذلك بعد الحصول على موافقة وزارة التربية والتعليم. وقد وزع الاختبار على معلمي المدرسة الواحدة، وتم جمع استجاباتهم في نهاية اليوم الدراسي وذلك لتعذر جمعهم في وقت واحد بسبب الحصص المطلوبة منهم.

4.3 المعالجات الإحصائية :

للإجابة على أسئلة الدراسة، تم إجراء التحليلات التالية:

1- حساب معاملات التمييز والصعوبة لكل فقرة من الفقرات.

2- حساب معاملات الثبات لمجالات الاختبار وللاختبار ككل.

3- استخراج الوسط الحسابي والانحراف المعياري لعلامات المعلمين على الاختبار، ومقارنة الوسط الحسابي مع المستوى المقبول للأداء وهو 64%.

4- حساب معامل الارتباط بين مجالات الاختبار، والدرجة الكلية كمؤشر على صدق البناء الداخلي للاختبار.

5- إجراء اختبار "ت" للعينات المستقلة للكشف عن الصدق التمييزي للاختبار.

الفصل الرابع

نتائج الدراسة

هدفت الدراسة الحالية إلى بناء اختبار لقياس كفايات المعلمين في مجال بناء اختبارات التحصيل، بحيث تتوافر لهذا الاختبار مؤشرات صدق وثبات وفاعلية فقرات، كما وسيتم الحديث عن مستوى أداء معلمي المرحلة الثانوية في الأردن على الاختبار.

ويمكن عرض النتائج وفقاً للترتيب التالي:

1) النتائج المتعلقة بالخصائص السيكومترية لهذا الاختبار.

2) النتائج المتعلقة بمستوى أداء معلمي المرحلة الثانوية على هذا الاختبار.

1.4 النتائج المتعلقة بالسؤال الأول :

يمكن عرض النتائج المتعلقة بالسؤال الأول وفقاً للترتيب التالي:

1.1.4 النتائج المتعلقة بالخصائص الإحصائية المميزة لفقرات الاختبار:

تم تحليل فقرات الاختبار للعينة التجريبية كما ورد في الفصل الثالث حيث تم حساب معاملات الصعوبة والتمييز لكل فقرة، وبعد الانتهاء من تطبيق الاختبار على أفراد العينة الاستطلاعية (الصدق والثبات)، وتصحيح الاختبار، حسبت معاملات الصعوبة والتمييز لكل فقرة.

ولحساب معامل الصعوبة لكل فقرة من فقرات الاختبار اعتمدت النسبة المئوية للمعلمين الذين أجابوا على الفقرة إجابة صحيحة عن الفقرة، أما بالنسبة للقدرة التمييزية للفقرة فقد حسبت من خلال إيجاد معامل ارتباط بين الفقرة والعلامة الكلية على الاختبار والجداول رقم (7)، (8) توضح قيم معاملات الصعوبة والتمييز لفقرات الاختبار.

جدول رقم (7)

معاملات الصعوبة لفقرات الاختبار لعينة الدراسة

رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة
1	0.72	24	0.57	47	*0.25
2	0.61	25	0.59	48	0.58
3	0.49	26	0.74	49	0.40
4	0.48	27	0.72	50	0.55
5	0.47	28	0.68	51	0.72
6	0.39	29	0.61	52	0.39
7	0.52	30	*0.90	53	0.35
8	0.5	31	0.53	54	0.68
9	0.42	32	0.47	55	0.51
10	0.38	33	0.56	56	0.90
11	0.54	34	0.4	57	0.65
12	0.33	35	0.46	58	0.56
13	0.3	36	0.58	59	0.40
14	0.36	37	0.45	60	0.49
15	0.68	38	0.43	61	0.56
16	0.48	39	*0.88	62	0.43
17	0.59	40	0.68	63	0.35
18	0.55	41	0.55	64	0.4
19	0.75	42	0.4	65	0.55
20	0.68	43	0.48	66	*0.85
21	0.61	44	0.40	67	0.3
22	0.36	45	0.39	68	0.68
23	0.62	46	*0.28	69	0.3
				70	0.4

* الفقرات ذات معاملات صعوبة متدنية أقل من 0.3 أو مرتفعة أكثر من0.75

جدول رقم (8)

قيم معاملات ارتباط الفقرة مع الدرجة الكلية (الدلالة التمييزية)

رقم الفقرة	معامل الارتباط	رقم الفقرة	معامل الارتباط	رقم الفقرة	معامل الارتباط
1	0.46	24	0.42	47	0.28
2	0.41	25	0.48	48	0.38
3	0.56	26	0.38	49	0.73
4	0.48	27	0.52	50	0.59
5	0.37	28	0.46	51	0.55
6	0.23	29	0.35	52	0.34
7	0.58	30	0.40	53	0.32
8	0.47	31	0.42	54	0.38
9	0.46	32	0.33	55	0.50
10	0.43	33	0.46	56	0.42
11	0.6	34	0.38	57	0.45
12	0.42	35	0.48	58	0.78
13	0.3	36	0.40	59	0.46
14	0.58	37	0.54	60	0.49
15	0.32	38	0.36	61	0.44
16	0.25	39	0.32	62	0.53
17	0.54	40	0.6	63	0.49
18	0.48	41	0.55	64	0.48
19	0.30	42	0.38	65	0.88
20	0.39	43	0.42	66	0.35
21	0.42	44	0.30	67	0.45
22	0.39	45	0.45	68	0.26
23	0.46	46	*0.16	69	0.39
				70	0.51

- الفقرات ذات قدرة تمييزية منخفضة (اقل من 0.2)

يتضح من الجدول رقم (7) ان معاملات الصعوبة لفقرات الاختبار تراوحت من 0.3 إلى 0.75 باستثناء الفقرة ذات الأرقام (30، 39,46، 47 ،66) التي كان معامل صعوبتهما على الترتيب (0.9، 0.88 ، 0.22 ، 0.28، 0.85) مما يشير

إلى أن غالبية فقرات الاختبار متوسطة الصعوبة، وأن هنالك فقرات سهلة (الفقرات: 30، 66) وأما بقية الفقرات فإن مستوى صعوبتها مقبول.

ويتضح من الجدول رقم(8) بأن القدرة التمييزية لفقرات الاختبار قد تراوحت من 0.2 إلى0.73 باستثناء الفقرة (46) حيث أن تمييزها (0.16) مما يدل على أن الاختبار يتمتع بقدرة تمييزه مقبولة.

2.1.4 النتائج المتعلقة بمؤشرات الصدق:

أن المؤشرات المستخدمة للكشف عن صدق الاختبار كانت على النحو التالي:

1) أولاً: صدق المحتوى:

تم تحديد ستة مجالات يمكن أن تغطي ما يحتاجه المعلم عند بنائه لاختبارات التحصيل التي يضعها بنفسه، ثم حلل كل مجال منها إلى أهداف مرحلية ثم تم تحليل الأهداف المرحلية النهائية إلى أهداف تفصيلية سلوكية وذلك باستخدام طريقة جانية (Gagne) في تحليل المهام وقد أمكن تغطية هذا المجال بعدد محدد من الفقرات ترتبط بهذه الأهداف وتم إختيار عينة من هذه الفقرات (70) موزعة على المهارات حسب أهميتها، وتمثل هذا المجال افضل تمثيلٍ، ولذلك يمكن الافتراض أن الاختبار صادق من حيث المحتوى.

2) ثانياً: الصدق التمييزي:

إلى جانب صدق المحتوى تم حساب مؤشر الصدق التمييزي للاختبار من خلال إجراء اختبار "ت" على عينتين مستقلتين، بلغ حجم كل منهما(40)فرداً، حيث قدم الاختبار في صورته النهائية إلى مجموعتين، إحداهما تلقت التعليم (طلبة الماجستير في القياس والتقويم والذين تلقوا تعليماً منتظما في بناء اختبارات التحصيل) والمجموعة الثانية لم تتلقى تعليما منتظماً شاملاً في مجالات الاختبار الستة وهم (مجموعة من المعلمين العاملين في وزارة التربية والتعليم ويمارسون وضع الاختبارات التحصيلية للعام الدراسي (2006/2005) ولم يتلقوا أي تعليم في هذا المجال. والجدول رقم (9) يبين نتائج ذلك.

جدول (9)

نتائج اختبار (ت) للمجموعتين المدربة وغير المدربة على الاختبار

المجموعة	المتوسط	الانحراف المعياري	قيمة(ت)	مستوى الدلالة
المجموعة المدربة	58.05	6.41	15.9	0.05
المجموعة غير المدربة	23.95	7.1		

من الجدول رقم (9) يتضح تفوق مجموعة طلبة الماجستير (المجموعة المدربة) على مجموعة المعلمين (المجموعة غير المدربة)، ولعل ذلك يعد مؤشراً على صدق الاختبار وقدرته على التمييز بين المجموعتين.

ثالثاً: صدق البناء الداخلي:

تم حساب معامل الارتباط بين مجالات الاختبار والدرجة الكلية كمؤشر على صدق البناء الداخلي للاختبار والجدول رقم (10) يبين نتائج ذلك.

جدول رقم(10)

معاملات الارتباط بين مجالات الاختبار والدرجة الكلية

المجال	معامل الارتباط
معرفة المعلم بالأهداف التعليمية.	0.80
معرفة المعلم بجدول المواصفات.	0.61
معرفة المعلم بكيفية انتقاء الطريقة المناسبة لقياس الأهداف.	0.73
معرفة المعلم بكيفية انتقاء الفقرات المناسبة.	0.88
معرفة المعلم بخصائص الاختبار الجيد.	0.81
معرفة المعلم بتصحيح الاختبار.	0.77
الدرجة الكلية.	0.80

من الجدول رقم (10) يتضح بان معاملات ارتباط المجالات مع الدرجة الكلية تراوحت ما بين (0.61-0.88) ، مما يشير إلى ان الاختبار يتمتع بدلالات صدق البناء الداخلي.

3.1.4 النتائج المتعلقة بمؤشرات الثبات:

وللتحقق من ثبات الاختبار تم استخدام الطرق التالية لتقدير ثبات الاختبار (التجزئة النصفية، كودر– ريتشاردسون– 20) والجدول رقم (11) يوضح ذلك.

جدول رقم (11)

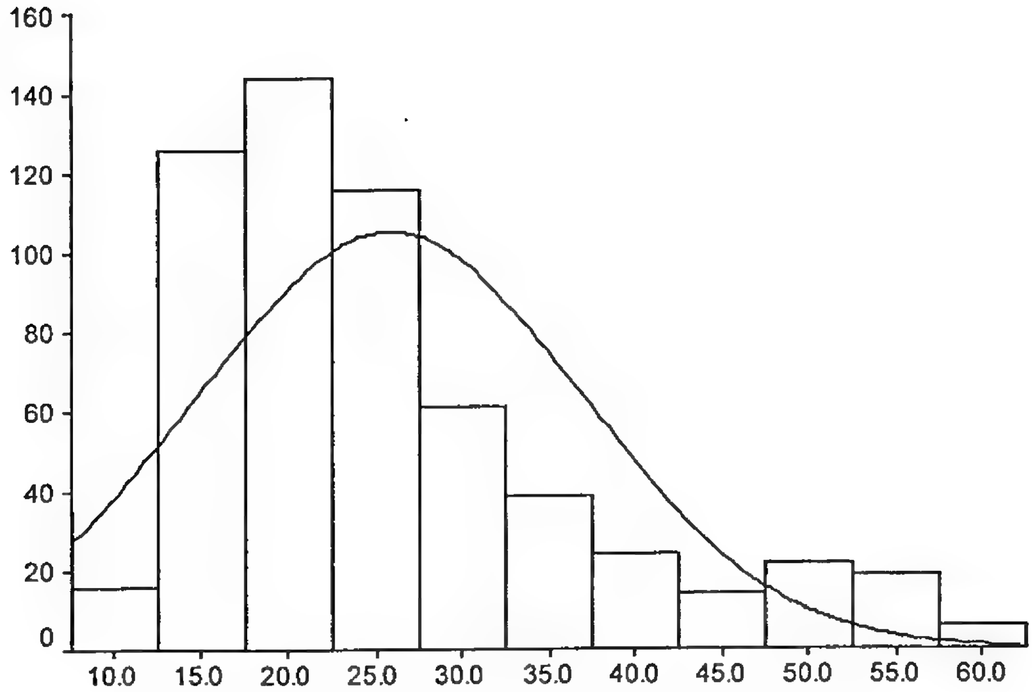
معامل ثبات الاتساق الداخلي لمجالات اختبار كفايات المعلمين في مجال بناء اختبارات التحصيل بالتجزئة النصفية و(Kr-20)

المجال	التجزئة النصفية	Kr-20
معرفة المعلم بالأهداف التعليمية	0.70	0.69
معرفة المعلم بجدول المواصفات	0.51	0.53
معرفة المعلم بكيفية انتقاء الطريقة المناسبة لقياس الأهداف	0.79	0.78
معرفة المعلم بكيفية انتقاء الفقرات المناسبة	0.79	074
معرفة المعلم بخصائص الاختبار الجيد	0.71	0.70
معرفة المعلم بتصحيح الاختبار	0.68	0.64
الدرجة الكلية	0.88	0.88

من الجدول رقم (11) يتضح بأن معاملات الثبات المحسوبة بطريقة التجزئة النصفية تراوحت ما بين 0.51 و 0.79 لمجالات الاختبار الستة، وأن معامل ثبات الاختبار ككل بلغ (0.88). أما معاملات الثبات المحسوبة بطريقة (Kr-20) تراوحت ما بين (0.53 و0.78) لمجالات الاختبار الستة، وأن معامل ثبات الاختبار ككل بلغ (0.88).

2.4 النتائج المتعلقة بالسؤال الثاني:

بعد تطبيق الاختبار على عينة الدراسة تم تصحيح الإجابات واستخرجت علامات جميع المعلمين على الاختبار وتم رسم مدرج تكراري، وذلك لعرض البيانات بصورة مختصرة والشكل رقم (5) يوضح ذلك.



شكل رقم (5)

المدرج التكراري لعلامات المعلمين على الاختبار

نلاحظ من هذا التوزيع أنه ملتو التواء موجباً، أي ان عدد كبير من المعلمين حصلوا على علامات منخفضة، بينما يقل هذا العدد كلما زادت العلامة.

كما تم حساب الوسط الحسابي والانحراف المعياري والمستوى المقبول للأداء ونسبة ما تم إتقانه لعلامات المعلمين على الاختبار ككل وعلى المجالات الستة للاختبار، حيث حسب المستوى المقبول و نسبة الإتقان وفقاً للعلاقات التالية

- درجة المستوى المقبول للأداء = عدد فقرات المجال × 64%
- نسبة ما تم إتقانه = (الوسط الحسابي للمعلمين على الاختبار أو المجال × 100) ÷ عدد فقرات الاختبار أو المجال). والجدول رقم (12) يوضح ذلك

جدول رقم (12)

الأوساط الحسابية والانحرافات المعيارية لعلامات المعلمين على المجالات الستة ودرجة المستوى المقبول ونسبة ما تم إتقانه.

المجال	عدد الفقرات	الوسط	الانحراف المعياري	درجة المستوى المقبول	نسبة ما تم إتقانه
معرفة المعلم بالاهداف التعليمية.	13	5.11	2.43	8.32	%39
معرفة المعلم بجدول المواصفات.	4	1.36	1.1	2.56	%34
معرفة المعلم بكيفية انتقاء الطريقة المناسبة لقياس الأهداف.	18	7.80	4.02	11.52	%43.3
معرفة المعلم بكيفية انتقاء الفقرات المناسبة.	11	3.2	1.7	7.04	%29
معرفة المعلم بخصائص الاختبار الجيد.	11	4.3	2.5	7.04	%39
معرفة المعلم بتصحيح الاختبار وتفسير نتائجه.	13	3.88	2.4	8.32	%30
الدرجة الكلية	70	26.60	11.51	45	%38

من الجدول السابق يتضح أن الوسط الحسابي لعلامات المعلمين على الاختبار يساوي (26.6) من أصل (70) أي ما يعادل (38 %) من هذه المهارات أي لم يصل إلى مستوى الأداء المقبول على هذا الاختبار وهو (64%) من هذه المهارات ويعادل في هذه الدراسة العلامة (45).

ويلاحظ أيضا أن متوسطات علامات المعلمين على المجالات الستة للاختبار لم تصل إلى مستوى الأداء المقبول.

الفصل الخامس
الخاتمة والمناقشة والتوصيات

1.5 الخاتمة:

هدفت الدراسة الحالية إلى بناء اختبار لقياس كفايات المعلمين في مجال بناء اختبارات التحصيل، حيث تم اشتقاق مؤشرات صدق وثبات الاختبار وكذلك الخصائص الإحصائية المميزة للفقرات، ومستوى الأداء المقبول على الاختبار (درجة القطع)، كما بحثت هذه الدراسة في مستوى معرفة معلمي المرحلة الثانوية في مجال بناء اختبارات التحصيل، ويتناول هذا الفصل مناقشة النتائج وتفسيرها كما يتناول بعض التوصيات.

2.5 مناقشة نتائج السؤال الأول:

1.2.5 النتائج المتعلقة بالخصائص الإحصائية للفقرة:

دلت النتائج المتعلقة بمعامل صعوبة الفقرات أنها تتراوح بين 0.3 و 0.75 عدا (5) فقرات. مما يشير إلى أن غالبية الفقرات متوسطة الصعوبة. أما بالنسبة لمعاملات التمييز لفقرات الاختبار فتراوحت بين 0.23 إلى 0.88 باستثناء فقرة واحدة، مما يدل على أن الاختبار يتمتع بقدرات تمييزية مقبولة.

ولعل وجود بعض الفقرات التي تقع خارج نطاق معامل الصعوبة أو التمييز المقبول قد يكون مناسباً من الناحية العملية حتى لو لم يكن مناسبا من الناحية الإحصائية، ولعل ما يبرر وجود مثل ذلك هو أنه تم تحليل مجالات الاختبار إلى أهداف نهائية، ثم حللت الأهداف النهائية إلى أهداف تفصيلية، وحيث أن هذه الأهداف محددة بدرجة عالية نسبياً، فإن فاعلية الفقرة في الاختبار المحكي تقاس بدرجة انسجامها مع الهدف.

2.2.5 النتائج المتعلقة بمؤشرات الثبات

دلت النتائج المتعلقة بالسؤال الأول أن الاختبار يتمتع بمستوى ثبات عالٍ نسبياً، إذ تراوحت معاملات ثبات الاتساق الداخلي بالطريقة النصفية لمجالات الاختبار ما بين 0.58 و 0.79، وأن معامل ثبات الاختبار ككل بلغ (0.88)، وبطريقة (Kr-20) تراوحت معاملات الثبات لمجالات الاختبار بين 0.59 و0.78 وأن معامل الثبات للاختبار ككل بلغ (0.88). وحيث أن اقل قيمة لمعامل الثبات قد كانت للمجال الثاني، فإن ذلك قد يعود ذلك إلى قلة عدد الفقرات في هذا المجال مقارنة مع بقية المجالات الأخرى، ونقصان عدد الفقرات يؤدي إلى نقصان التباين الحقيقي، وبالتالي نقصان معامل الثبات.

3.2.5 النتائج المتعلقة بمؤشرات الصدق:

أشارت النتائج المتعلقة بالصدق إلى أن الاختبار يتمتع بمؤشر صدق تمييزي عالٍ، حيث أظهرت نتائج إختبار (ت) لمجموعتين مستقلتين إحداهما تلقت التدريب في مجال بناء اختبارات التحصيل والأخرى لم تتلق مثل هذا التدريب، وجود فروق ذات دلالة إحصائية عند مستوى دلالة (P <= 0.05)، حيث كان الوسط الحسابي للمجموعة المدربة (58.05) والانحراف المعياري (6.41) أما بالنسبة للمجموعة غير المدربة فكان الوسط الحسابي (23.95) وانحراف معياري(7.1). مما يدل على أن الاختبار له القدرة على التمييز بين من تلقى تعليم في مجالات الاختبار وبين الذين لم يتلقوا تعليماً في هذا المجال.

كما أوضحت النتائج المتعلقة بمعامل الارتباط بين مجالات الاختبار والدرجة الكلية أما تتراوح بين (0.61–0.88) وجميعها دالة عند مستوى 0.001، مما يشير إلى ان الاختبار يتمتع بدلالات صدق البناء الداخلي.

3.5 مناقشة نتائج السؤال الثاني:

أظهرت النتائج المتعلقة بالسؤال الثاني وهو (ما مستوى أداء معلمي المرحلة الثانوية على الاختبار؟) أن مستوى معرفة معلمين المرحلة الثانوية لم يصل الى المستوى المقبول وهو (64%)، كما أن نسبة من أتقنوا هذه الكفايات من المعلمين متدنية جداً إذ بلغت (13 %) من المعلمين.

وتتفق نتائج هذه الدراسة مع نتائج دراسة كل من جرادات(1988)، الغمري (1998)التي أشارت إلى أن معرفة المعلمين وممارساتهم لكفايات بناء الاختبارات المدرسية لم تصل الى المستوى المقبول تربوياً.

وإذا كانت المعرفة شرط مسبق لتطوير اختبارات جيدة، فإن النتائج هنا تتفق مع دراسة العمري (1990) في الأردن التي توصلت إلى أن المعلمين لا يراعون معايير بناء الاختبار التحصيلي. إلا أن هذا يبقى مجرد افتراض، ولابد من إجراء دراسات تحليلية للاختبارات التي يطورها المعلمون في الأردن. إلا أن دراسة مارسو وبيج (Marso & Piggo,1992) التي حللت نتائج (225) دراسة في مجال بناء الاختبارات تؤكد صحة هذا الافتراض، فقد توصلت الى أن مهارات تطوير الاختبارات ليست بالمستوى المطلوب.

وقد يعود تدني مستوى أداء المعلمين على اختبار الكفايات في مجال بناء اختبارات التحصيل الى:

- عدم تعرض المعلمين أثناء الدراسة الى مساقات كافية في مجال القياس والتقويم وتطوير الاختبارات التحصيلية.
- عدم اهتمام المشرفين التربويين بتفحص أساليب القياس والتقويم –ومنها الاختبارات التحصيلية– لدى المعلمين.
- قلة الدورات التدريبية التي تعقد للمعلمين في هذا المجال، وإن وجدت فهي برامج نظرية وليست تدريبية عملية.
- عدم اهتمام المعلمين بتنمية قدراتهم والاطلاع على كل ما هو جديد في مجال تطوير الاختبارات المدرسية وانشغالهم بأولويات واهتمامات اخرى.

4.5 التوصيات:

من خلال النتائج التي تم التوصل إليها في هذه الدراسة، توصي الدراسة بما يلي:

1) لأن مستوى معرفة معلمي المرحلة الثانوية بكفايات تطوير الاختبارات التحصيلية متدنية ولم تصل إلى المستوى المقبول، فقد يكون من المناسب القيام بتنظيم دورات تدريبية لمعلمي المرحلة الثانوية لتزويدهم بالمعرفة النظرية والتطبيقية لتنمية مهاراتهم التعليمية في مجال الاختبارات التحصيلية.
2) اهتمام المشرفين التربويين بالاختبارات التحصيلية أثناء زياراتهم التقويمية للمعلمين وتركيزهم على تنمية مهاراتها لدى المعلمين.
3) إثراء برامج تدريب وتأهيل المعلمين قبل الخدمة بمساقات عن تطوير الاختبارات المدرسية
4) إجراء دراسات عن مستوى معرفة معلمي المرحلة الأساسية بكفايات بناء اختبارات التحصيل.
5) إجراء دراسات عن مستوى معرفة المشرفين التربويين بكفايات بناء اختبارات التحصيل.

المراجع

أ . المراجع العربية:

أبـو زيد، أمة الكريم طه. (1996). **الاحتياجات التدريبية لمعلمي العلوم في مرحلة التعلـيم الأساسـي بأمانـة العاصمة**. رسالة ماجستير غير منشورة،جامعة صنعاء، صنعاء: اليمن.

أبو سل، محمد عبدالكريم. (2002). **قياس وتقويم تعلم الطلبة**. عمان، الأردن: دار الفرقان.

أبـو لـبدة، سبع محمد. (1985). **مبادىء القياس النفسي والتقييم التربوي**. عمان الأردن: دار الفرقان.

ثـورندايك، روبـرت، هيجن، اليزابيث (1986). **القياس والتقويم في علم النفس والتـربية**. تـرجمة عـبدالله زيد الكيلاني، عبدالرحمن عدس، عمان:مركز الكتب الأردني.

جابـر، عبدالحمـيد. (1983). **الـتقويم التـربوي والقياس النفسي**. القاهرة: دار النهضة.

جـرادات، محمـد عـبد الرحمن. (1988). **مدى معرفة وممارسة معلمي العلوم للمرحلة الإعدادية لكفايات بناء الاختبارات المدرسية**، رسالة ماجستير غير منشورة، جامعة اليرموك، إربد: الأردن.

جـرونلاند، نـورمان. (1973). **إعـداد الاختـبارات محكـية المـرجع للتعلـيم الصفي**.ترجمة محمد الخوالدة. إربد- جامعة اليرموك.

الحميضـي، فاطمـة بنت عبد الله بن محمد. (1993). **تقويم الكفايات التعليمية في إعـداد وتصـحيح الاختبارات**. رسالة ماجستير غير منشورة. جامعة الملك سعود، كلية التربية: السعودية.

الخطيب، فريد مصطفى. (1993). **القياس والتقويم التربوي في المدارس**. عمان، الأردن: دار الهلال.

الـزيود، نـادر فهمي وعليان، هشام عامر. (1990). **مبادئ القياس والتقويم في التربية**. عمان، الأردن: دار الفكر للنشر والتوزيع.

ســـواقد، ساري. (1987). **عمل بنوك الأسئلة**، بحث غير منشور، جامعة اليرموك، إربد: الأردن.

الصراف، قاسم علي. (2002). **القياس والتقويم في التربية والتعليم**. الكويت: دار الكتاب الحديث.

عـــريفج، ســـامي ومصـــلح، خالد. (1987). **في القياس والتقييم**. عمان، الأردن: مطبعة رفيدي.

عطية، نعيم. (1970). **التقييم التربوي الهادف**. بيروت، لبنان: دار الكتاب اللبناني.

العمري، سمر امين جميل. (1991). **ممارسات التقييم الشائعة لدى المدرسين**. رسالة ماجستير غير منشورة، الجامعة الاردنية، عمان.

عـــلام، صـــلاح الدين محمود. (1985). بناء اختبار هدفي المرجع لقياس مهارات المعلمـــين فـــي تطوير الاختبارات المدرسية. **مجلة العلوم التربوية**. المجلد الثالث عشر، العدد الرابع، ص117-128.

عـــلام، صـــلاح الـــدين محمـــود. (1986). **تطورات معاصرة في القياس النفسي والتربوي**. كلية الآداب- جامعة الكويت. مطابع دار القبس.

عـــلام، صـــلاح الـــدين محمود. (1995). **الاختبارات التشخيصية مرجعية المحك المجالات التربوية والنفسية والتدريسية**. القاهرة: دار الفكر العربي.

عـــودة، أحمـــد سليمان. (1993). **القياس والتقويم في العملية التدريسية**. عمان، الأردن: دار الأمل.

عـــودة، احمد. (1990). تحليل نتائج الاختبارات المدرسية التي يعدها المعلمون بين الواقع والتطلعات. **مجلة جامعة مؤتة للبحوث والدراسات**، المجلد الخامس، العدد الأول، 247-269.

عـــودة، أحمـــد.(2002). **القـــياس والتقويم في العملية التدريسية**. الطبعة الثانية. الأردن:دار الأمل للنشر والتوزيع.

الغمـــري، صفية جبران مفرح. (1998). **مستوى معرفة معلمي المرحلة الأساسية في اليمن بمهارات وكفايات وتطوير الاختبارات المدرسية واتجاهاتهم نحو هــذه الاختــبارات. رسالة ماجستير غير منشورة**، جامعة اليرموك، إربد: الأردن.

القرشـــي، عــبد الفتاح. (1986). اتجاهات جديدة في أساليب تقويم الطلبة. **رسالة الخليج العربي**، العدد الثامن عشر، السنة السادسة، ص3-33.

المــنظمة العربية للتربية والثقافة والعلوم. (1974). **اجتماع خبراء نظم الامتحانات في بعض البلاد العربية**. الكويت: جامعة الدول العربية.

هويســت، ر. (1983). صــدق الاختــبارات المدرســية: دراسة في تحري صدق المحتوى. **المجلة العربية للبحوث التربوية**. المجلد الثالث، العدد الاول.

وزارة التربية والتعليم/إدارة التخطيط التربوي. (2005/2004). **التقرير الإحصائي**، عمان.

ب. المراجع الأجنبية:

Carey, Lou M. (1994). **Measuring and Evaluating school Learning.** Bouston: Allyn and Bacon.

Ebal, R. L. & Frisble, D. A. (1991). **Essentials of Educational Measurement.** New jersey: Primtic-Hall.

Gentry, Darrell. (1989). **Teacher-Made Test construction.** Paper presented at the annual meeting of the Mid-South Educational Research Association (Little Rock, AR, November 8-10, 1989), (ERIC Document Reproduction Service No. ED 313444).

Gronlund, N. E. (1977). **Constructing Achievement Tes.** London: Print ice-Hall INC. chapters 2-5.

Hopkins, C. D. and Antes, R.L (1978) .**Classroom Measurement and Evaluation.** Itasca: Peacock Publishing INC.

Kerber, P. E. (1976). **A study of Iowa sixth grade teacher's knowledge of tests and measurement concepts.** Dissertation Abstract, 36:8, P5220. University of Iowa.

Marso, Ronald N. and pigge, Fred L. (1988). **An Analysis of teacher-made tests:testing practices, cognitive demands, and item construction errors.**(Paper presented at the Annual Meeting of the National council on measurement in education. (New Orleans, April 6-8).

Marso, Ronald N. and pigge, Fred L. (1992). **Classroom Teachers Knowledge and Skills Related to the Development and use of Teacher-Made Tests.** Paper presented at the annual meeting of the American Educational Research Association (ERIC Document Reproduction Service No. ED 346148).

Miller, M. Ann, H. (2002). **The Effects of Teaching content literacy strategies Upon a Conveniently Selected Sample of at-Risk 9-12Th Grade Students.** Unpublished. Phd Dissertaion.Loyola university of Chicago(USA).

Newman, D. C. (1981). **Teacher competency in classroom tes-ting practices.** Dissertation Abstracts 42:31 unpublished Doctoral dissertation university of Georgia.

Oosstrhof, Albert. (1994). **Classroom Application of Educ-ational Measurement.** Second Edition, New York, Macmillan Publishing Company.

Stiggins, R. J. & Conklin, K. F. (1988). **Teacher Training In Assessment Washington**, Dc: Office of Educational Research and Improvement, US Department of Education. In Al-Dosary, Rashid Hammed (1993). Assessing The Competency of Teachers, Curriculum Specialists, and Prospective Teachers In Educational Measurement In Bahrain. Doctoral Dissertation Unpublished. Michigan State University. Educational Psychology and Special Education.

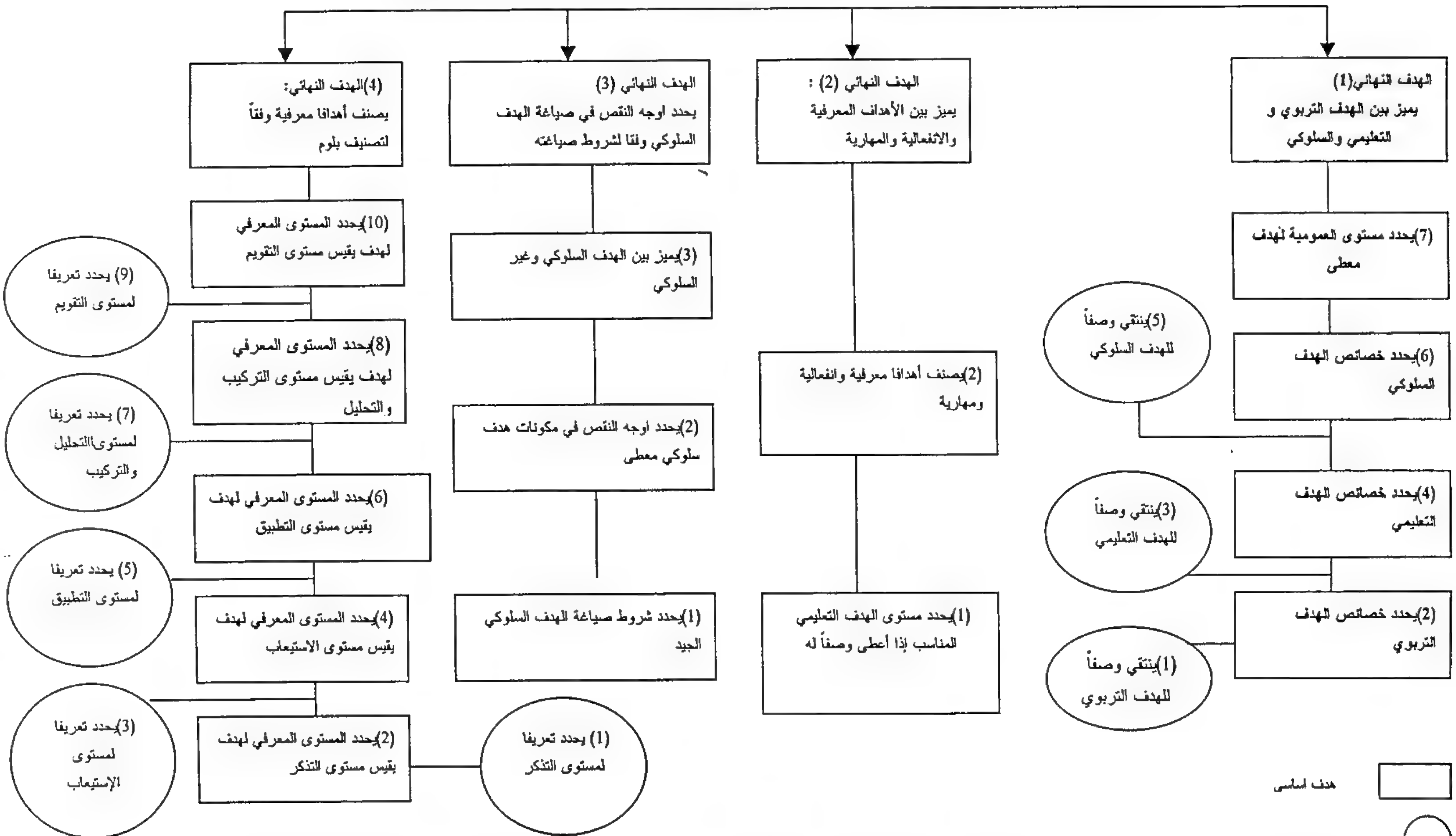
Tyler, R. (1988). **Evaluation for Utilization**. In keeves (Ed.), Educational Research, Methodology and Measurement An International Hand book. Pergamon Press. New York. (155-164).

Williams, Jane M. (1991). **Writing Quality Teacher-made Tests A Handbook for teachers**. (ERIC Document Reproduction Service No. ED 349726).

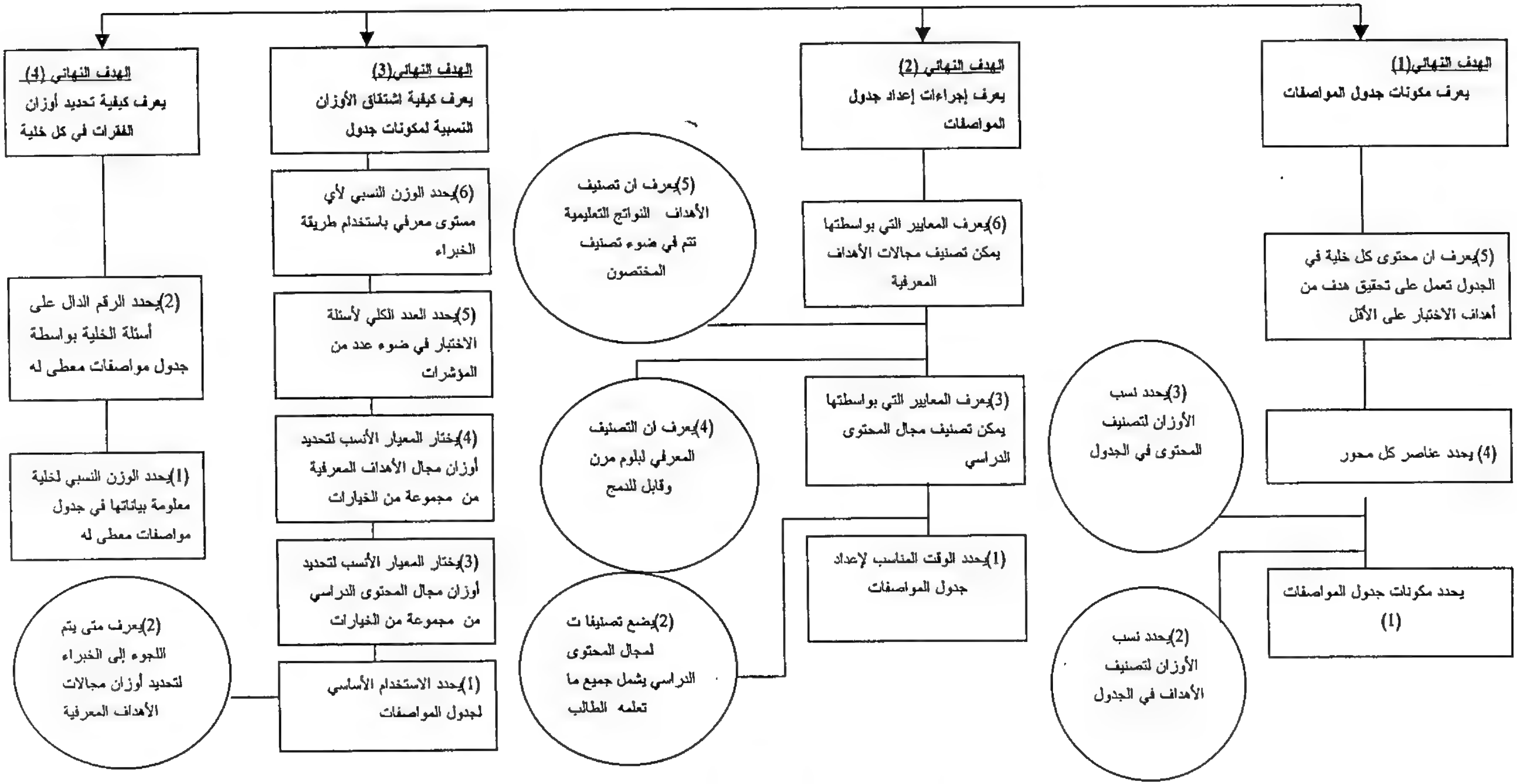
الملحق(أ)

تحليل ألاهداف المرحلية النهائية إلى أهداف تفصيلية سلوكية

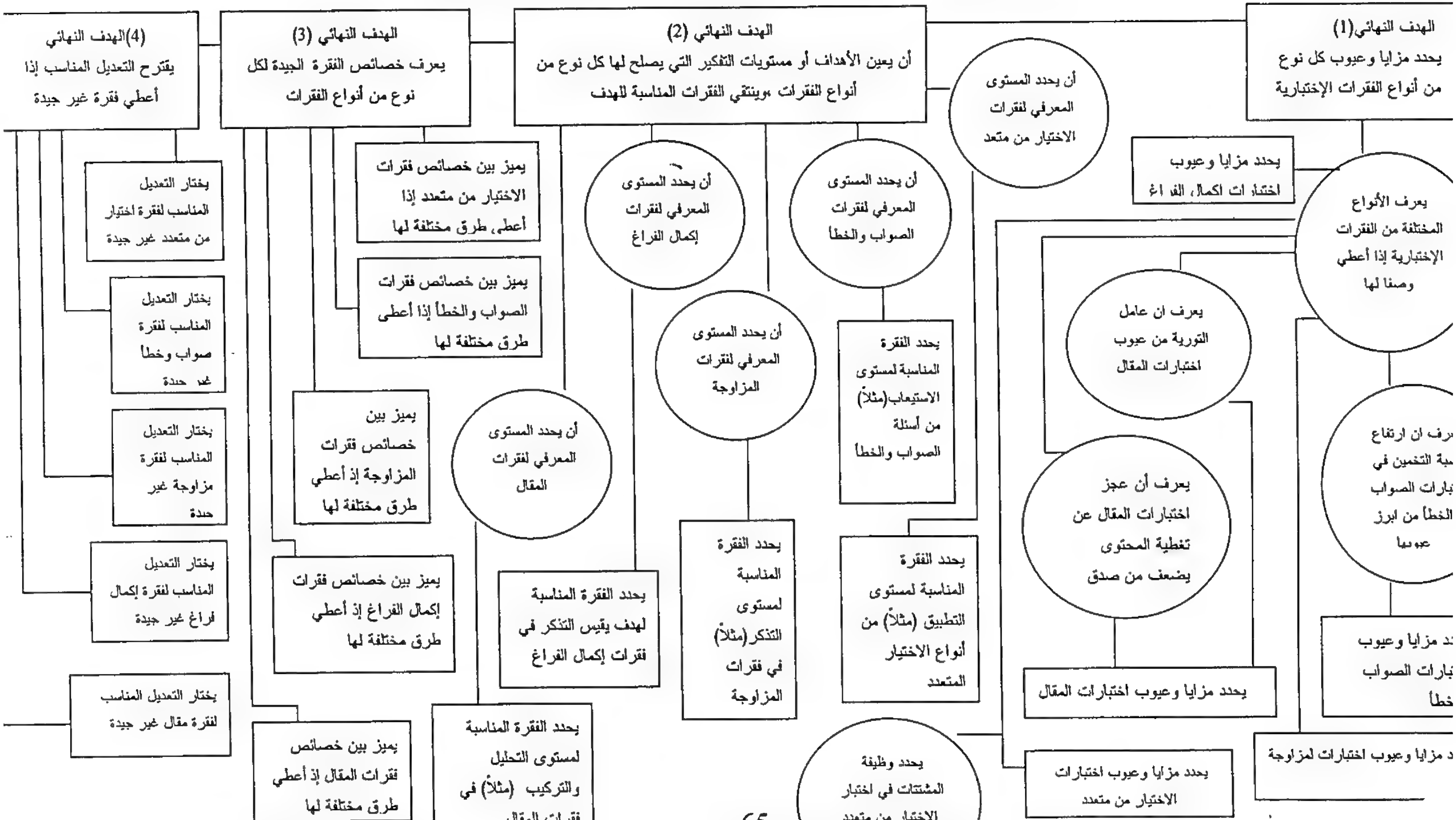
مخطط(1) الأهداف التعليمية ودورها في عملية التعلم والتعليم وفقاً لأصنافها:



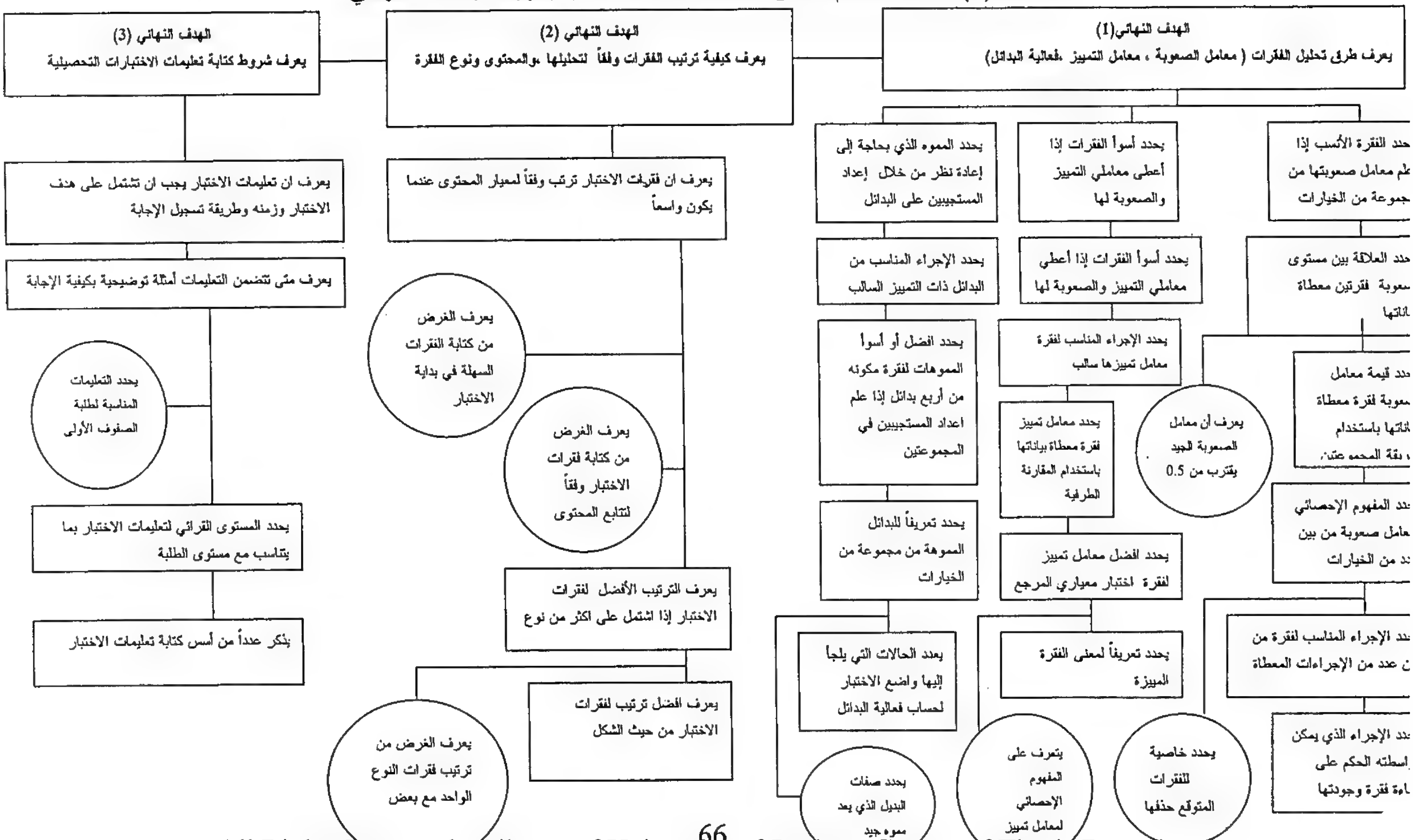
مخطط(2) جدول المواصفات ودوره في بناء اختبارات التحصيل



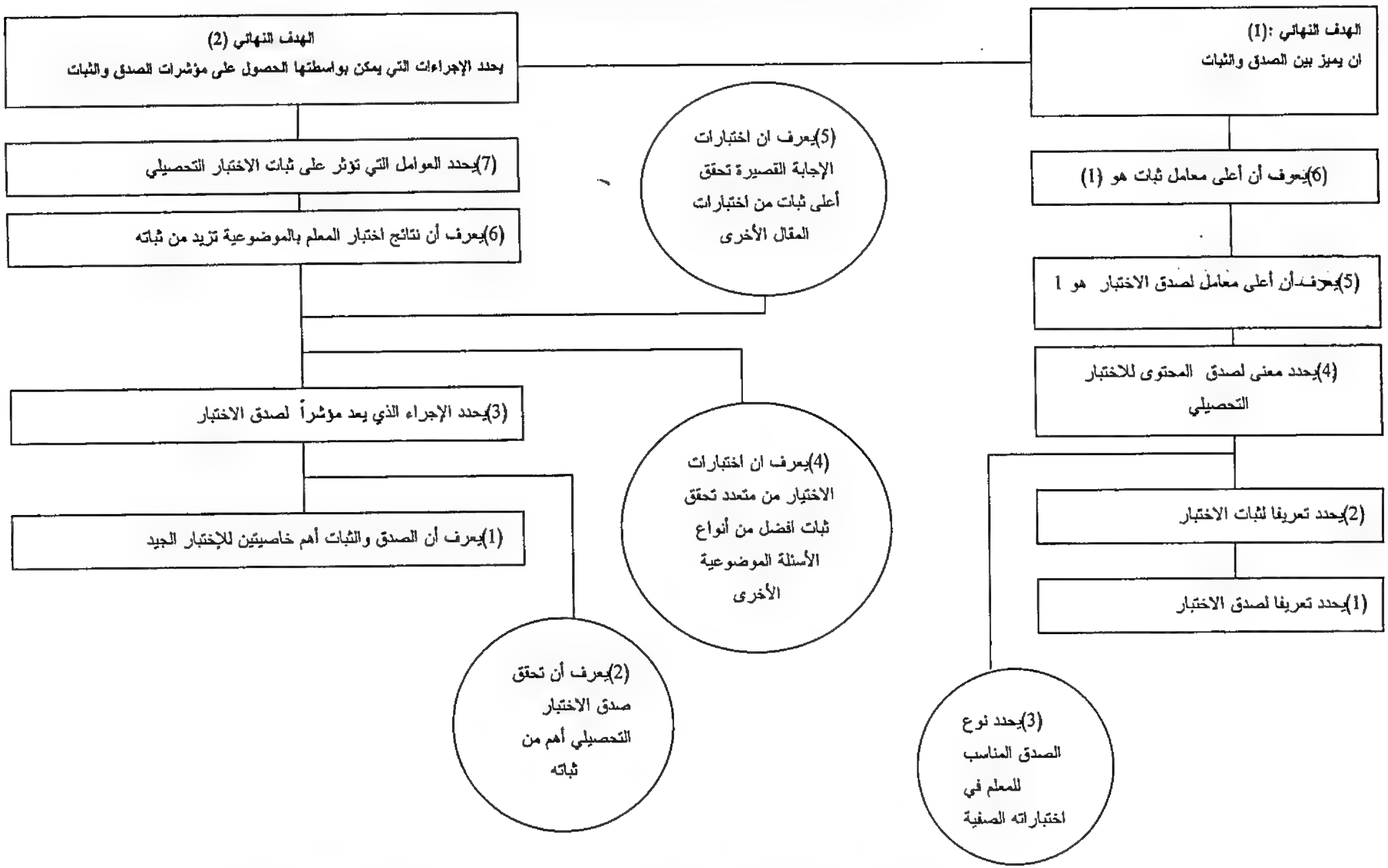
مخطط(3) معرفة المعلم بكيفية اختيار الفقرات المناسبة لقياس أهدافه



مخطط (4) معرفة المعلم بطرق انتقاء فقرات الاختبار وتهيئته بشكله النهائي



مخطط (5) معرفة المعلم بخصائص الاختبار الجيد



المجال الاول	مناسبة المجال لهدف الاختبار			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
1)معرفة المعلم للاهداف التعليمية ودورها في عملية التعلم والتعليم وفقاً لاصنافها						

الاهداف النهائية للمجال الاول	مناسبة الاهداف النهائية للمجال			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
أ) ان يميز المعلم بين الهدف التربوي والتعليمي والسلوكي .						
ب)ان يميز المعلم بين الاهداف المعرفية والانفعالية والمهارية.						
ج) ان يحدد اوجه النقص في صياغة الهدف السلوكي ، وفقاً لشروط صياغته .						
د)ان يصنف اهدافاً معرفية سلوكية وفقا لتصنيف بلوم						

اذا كنت ترى بأن هنالك اهدافا لم يشملها المجال فضلاً اذكرها

....

....

....

....

....

المجال الثاني	مناسبة المجال لهدف الاختبار			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
2) معرفة المعلم بجدول الموصفات ودوره في بناء اختبارات التحصيل .						

الاهداف النهائية للمجال الثاني	مناسبة الاهداف النهائية للمجال			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
أ) ان يعرف مكونات جدول المواصفات .						
ب)ان يعرف إجراءات اعداد جدول المواصفات						
ج) ان يعرف كيفية اشتقاق الاوزان النسبية لمكونات جدول المواصفات (المحتوى ، او الاهداف ، المستويات المعرفية ، عدد الفقرات)						
د) ان يعرف كيفية تحديد اعداد الفقرات في كل خلية						

اذا كنت ترى بأن هنالك اهدافا لم يشملها المجال فضلاً اذكرها

....

....

....

....

....

المجال الثالث	مناسبة المجال لهدف الاختبار			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
3) معرفة المعلم بكيفية اختيار الفقرات المناسبة لقياس اهدافه .						

الاهداف النهائية للمجال الثالث	مناسبة الاهداف النهائية للمجال			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
أ) ان يحدد مزايا كل نوع من انواع الفقرات الاختبارية.						
ب) ان يحدد عيوب كل نوع من انواع الفقرات الاختبارية						
ج)ان يعين الاهداف المعرفية او مستويات التفكير التي يصلح لها كل نوع من انواع الفقرات ، وينتقي الفقرات المناسبة للهدف						
د) ان يعرف خصائص الفقرة الجيدة لكل نوع من انواع الفقرات الاختبارية .						
هـ) ان يختار التعديل المناسب اذا اعطيت له فقرة غير جيدة .						

اذا كنت ترى بأن هنالك اهدافا لم يشملها المجال فضلاً اذكرها

....

....

....

المجال الرابع	مناسبة المجال لهدف الاختبار			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
4) معرفة المعلم بطرق انتقاء فقرات الاختبار وتهيئة بشكله النهائي						

الاهداف النهائية للمجال الرابع	مناسبة الاهداف النهائية للمجال			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
أ) ان يعرف طرق تحليل الفقرات (معامل الصعوبة ، معامل التمييز ، فعالية البدائل)						
ب)ان يعرف كيفية ترتيب الفقرات وفقاً لتحليل الفقرة أو محتواها أو نوعها .						
ج) ان يعرف شروط كتابة تعليمات الاختبارات التحصيلية.						

اذا كنت ترى بأن هنالك اهدافا لم يشملها المجال فضلاً اذكرها

....

....

....

....

....

المجال الخامس	مناسبة المجال لهدف الاختبار			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
5)معرفة المعلم بخصائص الاختبار الجيد .						

الاهداف النهائية للمجال الخامس	مناسبة الاهداف النهائية للمجال			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
أ) ان يميز بين مفهوم الصدق والثبات .						
ب) ان يحدد الاجراءات التي يمكن بواسطتها الحصول على مؤشرات عن الصدق.						
ج) ان يحدد الاجراءات التي يمكن بواسطتها الحصول على مؤشرات عن الثبات.						

اذا كنت ترى بأن هنالك اهدافا لم يشملها المجال فضلاً اذكرها

....

....

....

....

....

المجال السادس	مناسبة المجال لهدف الاختبار			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
6) معرفة المعلم بتصحيح الاختبار وتفسير نتائجه .						

الاهداف النهائية للمجال السادس	مناسبة الاهداف النهائية للمجال			الصياغة اللغوية		
	مناسب	يحذف	التعديل المقترح	سليمة	غير سليمة	التعديل المقترح
أ)ان يعرف طرق تصحيح اختبارات التحصيل بنوعيها المقالية والموضوعية.						
ب) ان يتعرف على الفرق بين الطريقة الكلية والطريقة التحليليةفي تصحيح الاسئلة المقالية .						
ج) ان يقدم وصفاً للبيانات من خلال مقاييس النزعة المركزية والتشتت والمئينات والدرجات المعيارية .						
د)ان يفسر الدرجة الخام بالطرق الاحصائية						

اذا كنت ترى بأن هنالك اهدافا لم يشملها المجال فضلاً اذكرها

....

....

....

....

....

بسم الله الرحمن الرحيم

أخي المعلم /أختي المعلمة،،،،

السلام عليكم ورحمة الله وبعد

تلعب الإختبارات دوراً مهماً في العملية التربوية ، فعلى ضوء نتائجها يتم اتخاذ الكثير من القرارات التي تنعكس آثارها على الطالب بشكل خاص ، وكل الأطراف المشاركة في العملية التربوية بشكل عام ، ولذلك يجري الباحث **دراسة تهدف إلى تطوير اختبار لقياس كفايات المعلمين في مجال بناء اختبارات التحصيل** . فنرجوا منك التعاون معنا لما لذلك من أهمية في معرفة كفايات المعلم الأردني في هذا المجال . مع العلم ان نتائج هذه الدراسة لا تستخدم إلاّ لأغراض البحث العلمي ولا داعي لكتابة اسمك .

القسم الأول : معلومات عامة:

ضع إشارة × في الفراغ الذي يسبق الإجابة التي تنطبق عليك.

الجنس : ☐ ذكر ☐ أنثى

الخبرة: ☐ خمس سنوات فأقل ☐ اكثر من خمس سنوات

المؤهل العلمي : ☐ دكتوراه ☐ ماجستير

☐ بكالوريوس +دبلوم ☐ بكالوريوس

☐ دبلوم كلية مجتمع

القسم الثاني : اختبار كفايات بناء اختبارات التحصيل.

يتكون هذا الاختبار من(70) فقرة من نوع الاختيار من متعدد . وكل فقرة لها أربعة بدائل . لذا ارجوا ان تستجيب لها جميعاً. ويرجى قراءة كل فقرة بعناية ثم وضع إشارة (×) على ورقة الإجابة المرفقة بهذا الاختبار في المربع البديل الذي تعتبره صحيحاً. كما هو موضح بالمثال التالي:

البديل / الفقرة	أ	ب	جـ	د
1		×		
2	×			

فقرات الاختبار

1) يشير الفعل في الهدف السلوكي إلى ....

أ) ما يؤديه المعلم في الصف.

ب) ما يتوقع من الطالب القيام به.

ج) طريقة القياس ومستواه .

د) ما يجب تضمينه في المنهج.

2) أي الشروط التالية لا تدخل ضمن مواصفات الهدف السلوكي الإجرائي ؟

أ) ان يكون عاماً .

ب) انه يمكن ملاحظته.

ج) انه يحتوي على الحد الأدنى للقياس .

د) انه يوضع على أساس نشاط التلميذ.

3) تعبر الأهداف التربوية بشكل عام عن ....

أ) مجموعة من المبادئ العامة التي تعبر عن طموحات وآمال المجتمع

ب) الأهداف السلوكية الخاصة بكل درس

ج) الأهداف الإجرائية التي يحاول المعلم تحقيقها داخل الفصل

د) الأهداف التعليمية التي يمكن تحقيقها من خلال مادة دراسية واحدة

4) واحدة من العبارات التالية **لا تشير** إلى مكونات الهدف التعليمي ؟

أ) السلوك النهائي من جانب التلميذ

ب) الظرف أو الشرط الذي سيؤدي في ظله المتعلم السلوك

ج) المتطلبات السابقة لتحقيق الهدف التعليمي

د) معيار الأداء

5) " ان يظهر الطالب التزاماً بخدمة الآخرين عن طريق العضوية في أحد تنظيمات الخدمة العامة داخل المدرسة "

يصنف الهدف السابق من مستوى الاهداف تحت المجال ....

أ) المعرفية

ب) الانفعالية

ج) الحس حركية

د) لاشيء مما ذكر

6) العبارة الهدفية والتي تنتمي إلى المجال " النفس حركي " هي ....

أ) ان يطبع الطالب أربعين كلمة في الدقيقة دون ارتكاب اكثر من خمسة أخطاء في اختبار لا تزيد مدته عن خمس دقائق

ب) ان يشارك الطالب باستمرار في حملات جمع التبرعات للمنكوبين في العالم

ج) ان يعيد الطالب كتابة معادلة رياضية بالكلمات بدلاً من الرموز

د) ان يسمي الطالب مخترع الجرس الكهربائي

7)الفعل الذي **لا يعبر** عن سلوك قابل للملاحظة والقياس عند صياغة الأهداف التعليمية الدقيقة فيما يلي هو ....

أ) ان يدرك

ب)ان يرسم

ج) ان يذكر

د) ان يعين على الخريطة

8) النقص في صياغة الهدف " ان يظهر الطالب قدرة في صياغة الأهداف التعليمية الدقيقة بتحصيل علامة 80% على الأقل في اختبار إجباري من نوع اختيار من متعدد عن الأهداف السلوكية " هو ......

أ) الافتقار إلى السلوك القابل للملاحظة .

ب) الافتقار إلى الحد الأدنى للمعيار/مستوى الأداء

ج) الافتقار إلى (أ ، ب)

د) اختلاف السلوك الختامي المقصود عن السلوك المطلوب بيانه أو الدلالة عليه .

9) يشترط في الهدف التعليمي الدقيق ان يتضمن الاداء الذي....

أ) يتوقع من الطالب إنجازه بعد عملية التدريس .

ب) يتوقع ان يقوم به المعلم قبل البدء بالتدريس .

ج) يقوم به الطالب أثناء عملية التدريس .

د) يتوقع ان يقوم به المعلم والطالب أثناء عملية التدريس .

10) العبارة الهدفية التي تدل على مستوى التطبيق فيما يلي هي ان :....

أ) يستخدم الطالب فكرة تجريدية عندما يتم تحديد كيفية استخدامها .

ب) يستخدم الطالب الفكرة التجريدية بطريقة صحيحة إذا توافرت له الظروف الملائمة والتي لا يحدد فيها كيفية الاستخدام .

ج) يعيد كتابة معادلة في الرياضيات بالكلمات بدل الرموز .

د) ان يكتب محضراً للاجتماع .

11)يصنف الهدف التالي " أن يسمي الأجزاء المرقمة في الرسم " بمستوى.....

أ) الفهم . ب) المعرفة . ج) التطبيق . د) التحليل .

12) يصنف الهدف التالي " يترجم قطعة نثرية من اللغة العربية الى اللغة الإنجليزية " بمستوى

أ) التذكر ب) الفهم ج)التطبيق د) التحليل

13) **ما مقترحاتك من اجل ان تستعيد جامعة الدول العربية قوتها في الوقت الحاضر ؟**

*المستوى المعرفي الذي يقيسه السؤال السابق هو ....

أ) تقويم ب) تركيب ج) تحليل د)تذكر

14) الفائدة الرئيسية من استخدام جدول المواصفات في بناء أدوات القياس هي

أ) توفير الوقت

ب) تجعل فقرات الاختبار اكثر سهولة

ج) تزيد موضوعية الاختبار

د) تحسن في تمثيل المجال السلوكي

15) يتكون جدول المواصفات من بعدين هما ....

أ) تقديرات ومعايير

ب) أهداف ومحتوى

ج) احتياجات المعلم والمتعلم

د)أهداف و أنشطة

16) إذا كان وزن المحتوى 20% ووزن الهدف المعرفي (الاستيعاب)=30% وكان عدد أسئلة الاختبار 50 فقرة فإن عدد الأسئلة للخلية الناتجة غن تقاطع النسبتين في جدول المواصفات هي

أ) 2 ب) 3 ج) 5 د) 6

17) يرجع الوزن النسبي للمحتوى الدراسي داخل جدول المواصفات إلى جميع المحكات التالية <u>ما عدا</u> واحد هو :.......

أ) أهمية الموضوع كمتطلب سابق للتعلم اللاحق

ب) عدد الصفحات المخصصة للموضوع

ج) الزمن المخصص لتدريس كل موضوع من الموضوعات

د) عدد الجداول والأشكال المذكورة

18) أي الإختبارات التالية يتوقع أن يكون اكثر صدقاً وثباتاً ؟ ....

أ) الشفوية ب) المقالية ج) الموضوعية د) التحريرية

19) أهم ما تتميز به الاختبارات الموضوعية عن اختبارات المقال ....

أ) توفير وقت المعلم اثناء عملية الاعداد

ب) قياسها مستويات مختلفة من الأهداف .

ج) اشتمالها على عينة اكثر تمثيلاً لمحتوى المادة الدراسية .

د) قلة الوقت اللازم للإجابة عنها .

20) يمكنك ان تحقق ثبات الاختبار المقالي بأن : ....

أ) تزيد عدد الأسئلة في الاختبار

ب) تقتصر التصحيح على مصحح واحد

ج) تنظيم أسئلة طويلة تستغرق وقتاً طويلاً للإجابة

د) تقلل عدد الأسئلة في الاختبار

21) احدى الفقرات التالية أفضل ما يمكن كمثال على فقرات(الصح- الخطأ)وهي:

أ) كتب طه حسين ألايام وعبقرية عمر.

ب) بويع شوقي اميراً للشعراء وعين وزيراً عام 1940م.

ج) تنتهي جملة(الصح- خطأ) دائماً بفترة زمنية مع علامة استفهام.

د) كان الكتاتيب يجرون الامتحانات شفوياً.

22) تشير جميع الجمل الآتية إلى مميزات اختبار الخطأ والصواب **ما عدا** جملة واحدة هي :

أ) سهولة الاعداد وقدرتها الهائلة على تغطية أجزاء كبيرة من المقرر

ب) ارتفاع اثر التخمين

ج) سهولة التصحيح

د) تحتاج إلى وقت قصير للإجابة

23) أي من البنود التالية يتفق مع قواعد صياغة السؤال الجيد من نوع الاختيار من متعدد .....

أ) تضمين الكلمات المكررة في البدائل

ب) استخدام عبارات من نوع (كل ما سبق)

ج) تضمين المشكلة في رأس السؤال (المتن)

د) استخدام أدوات النفي في رأس السؤال

24) لكي نتجنب التلميح بالإجابة بين المشتتات في الإختيار من متعدد ، نضع الإجابة الصحيحة في .....

أ) نهاية المشتتات

ب) مكان غير ثابت

ج) مكان ثابت

د) بداية المشتتات

25) تشير جميع العبارات التالية إلى القواعد التي يجب مراعاتها ليتصف الإختبار بالموضوعية **ما عدا** عبارة واحدة هي : ....

أ) وضوح الأسئلة وخلوها من الغموض

ب) شمول الأسئلة لمختلف أجزاء المادة

ج) تصميم الأسئلة على شكل أسئلة مقالية طويلة

د) جعل لغة الأسئلة سهلة في مستوى التلاميذ .

26) تمتاز الاختبارات الموضوعية بــ.......

أ)قياس قدرة التلاميذ على الابتكار والتنظيم

ب) اختصار الزمن بالنسبة لواضع الأسئلة .

ج) الصدق والثبات والموضوعية

د)إعطاء التلاميذ فرصاً لشرح وإيضاح المفاهيم المختلفة

27) أي أنواع الفقرات التالية تتأثر إجابات الطلاب عنها بدرجة اكبر بعامل التخمين؟

أ) التكملة

ب) الصواب والخطاء

ج)الاختيار من متعدد

د)المزاوجة

28) " طوكيو أهم مدينة في اليابان " تعد فقرة صواب أو خطأ غير جيدة لأنها ...

أ) غاية السهولة

ب)مختصرة جداً

ج)تقيس تذكر الحقائق

د)تحتمل اكثر من إجابة واحدة

29) تأمل الفقرة التالية ثم أجب على السؤال الذي يليها

الذي يأكل اللحوم من بين التالية هو ......
أ) الزيتون ب) الأسد ج) الحمار د) الدجاجة

عيب هذه الفقرة هو في ....

أ) مقدمتها ب) البديل الصحيح ج) المموهات د) مقدمتها وبدائلها

30) تأمل الفقرة التاليه ثم اجب على السؤال الذي يليها

وفق بين الدولة في العامود (أ) وعاصمتها في العامود (ب)

الدولة	العاصمة
( ) الأردن	القاهرة -1
( ) السعودية	عمان -2
( ) مصر	الرياض -3

أهم عيب ان وجد في الفقرة السابقة هو

أ) عدم صلاحية بعض الإجابات التي في العامود الثاني

ب) تحتوي على محدد نوعي

ج)عدد العناصر في عمود الإجابات مساو لعدد العناصر في عمود المقدمات

د)المفردة صالحة كما هي مكتوبة

31) أي من البنود التالية يتفق وقواعد صياغة السؤال الجيد من نوع الصواب والخطأ ؟

أ) كتابة كلمات تدل على التعميم مثل (دائماً ، أبداً ، غالباً)

ب) الإكثار من عبارات الصواب والخطأ ضمن الإختبار الواحد

ج) نقل عبارات بنص الكتاب

د) تساوي طول العبارات الصواب والعبارات الخطأ

32) أي من البنود التالية يتفق مع قواعد صياغة السؤال الجيد من نوع المزاوجة ؟

أ) ان يتراوح عدد عبارات قائمة الاستجابات بين 10 –15 عبارة

ب) تساوي قائمة المثيرات مع قائمة الاستجابات في عدد العبارات

ج) تجانس كل من قائمة المثيرات وقائمة الاستجابات

د) تنقل قائمة المثيرات بنص الكتاب

33) أي من البنود التالية يتفق مع قواعد صياغة السؤال الجيد من نوع الإكمال ؟

أ) الفراغات في بداية العبارة

ب) الفراغات في نهاية أو وسط العبارة

ج)نقل العبارات بنصها من الكتاب

د) شمول العبارة على فراغات كثيرة

34) أي من نواتج التعلم التالية تقيسها أسئلة الصواب والخطأ ؟
أ) تنظيم الأفكار وترتيبها
ب) ابتكار أفكار جديدة
ج) التعبير الكتابي
د) تفسير العلاقة بين السبب والنتيجة

35) أي من نواتج التعلم التالية تقيسها أسئلة المزاوجة ؟
أ) معرفة الحقائق والمفاهيم
ب) ابتكار أفكار جديدة
ج) تحليل وتفسير المعلومات
د) التعبير الكتابي

36) أي من نواتج التعلم التالية تقيسها أسئلة الإكمال ؟
أ) تنظيم الأفكار وترتيبها
ب) استدعاء الحقائق والمفاهيم
ج) ابتكار أفكار جديدة
د) نقد وتحليل الأحداث

37) أي من الأسئلة التالية من نوع الإكمال تتفق وقواعد صياغة السؤال الجيد
أ) ......فقدت قوتها عقب تحطيم الأسطول في أبى فير البحري
ب) تعد الحملة الفرنسية على ....
ج) ....، ....، ....من نتائج الحملة الفرنسية
د) حاولت إنجلترا الاستيلاءعلى مصر فأرسلت حملة عسكرية بقيادة فريزر عام ...

38) في الاختبارات التي يكتبها المعلم لتلاميذه ينبغي ترتيب الفقرات
أ) من السهل إلى الصعب بدءاً بمفردات الصواب والخطاء وانتهاء بأسئلة المقال
ب) من السهل إلى الصعب بدءاً بأسئلة المقال وانتهاء بمفردات الصواب والخطأ
ج) من الصعب إلى السهل بدءاً بأسئلة المقال وانتهاء بمفردات الصواب والخطأ
د)بطريقة عشوائية

*تأمل البيانات المتعلقة بفقرة من نوع الاختيار من متعدد ،في الجدول التالي ثم اجب عن الأسئلة من(39) إلى(43) :

البدائل	المجموعة العليا (50)	المجموعة الدنيا (50)
أ	18	10
ب	2	20
ج ✓	30	20
د	0	0

39)معامل تميز الفقرة يساوي .....

أ) 0.2 ب)0.3 ج) 0.5 د) 0.7

40) معامل صعوبة الفقرة يساوي .....

أ) 0.2 ب) 0.6 ج) 0.5 د) 1

41) المموه الذي ينبغي إعادة النظر فيه هو .....

أ) أ ب) ب ج) ج د) د

42) المموه الضعيف الذي ينبغي حذفه واستبداله هو ......

أ) أ ب) ب ج) ج د) د

43) افضل المموهات فعالية هو ....

أ) أ ب) ب ج) ج د) د

---

44) في الاختبارات معيارية المرجع يفضل الفقرات ذات معامل صعوبة قيمته

أ) 0.80 ب) 0.50 ج) 0.20 د) 0.25

45) الفقرة التي ينبغي استبعادها من الاختبار معياري المرجع هي الفقرة التي معامل تمييزها

أ) .25 ب) .95 ج) - .90 د) 1

*. تأمل البيانات المتعلقة بفقرة اختبارية صياغية والتي صححت من (5) نقاط ، ثم أجب عن الأسئلة 46 – 47:

مجموعة الأداء المرتفع	5، 5 ، 2، 4، 1، 3	
مجموعة الاداء المنخفض	2، صفر، 5، صفر، 1، 2	

46) معامل صعوبة هذه الفقرة يساوي -----.

أ) 0.166 ب) 0.33 ج) 1 د) 0.50

47)معامل تمييز هذه الفقرة يساوي -----.

أ) 0.166 ب) 0.33 ج) 1 د) 0.50

48) أفضل الفقرات في الاختبارات معيارية المرجع هي التي -----.

أ) تمييزها = 0.5 وصعوبتها = 0.5

ب) تمييزها = -1 و صعوبتها = 1

ج) تمييزها = 1 وصعوبتها = 0.5

د) تمييزها = 0.3 و صعوبتها = 0.7

49) يعد الاختبار صادقاً إذا كان ....

أ) تحصيلياً ،ويقيس قدرات عقلية

ب) يقيس مقدار الخاصية التي صمم لقياسها

ج)يقيس أهدافا تربوية عامة

د)جميع ما سبق ذكره

50) قدرة الاختبار على التمييز تعني ......

أ)النسبة المئوية لعدد المتعلمين الذين أجابوا بشكل صحيح على كل سؤال في الاختبار

ب)قدرة كل سؤال على التمييز بين التلاميذ الذين حصلوا على درجات عالية ومنخفضة

ج) قدرة كل سؤال من الأسئلة على التمييز بين التلاميذ الذين حصلوا على درجات عالية

د) قدرة كل سؤال من الأسئلة على التمييز بين التلاميذ الذين حصلوا على درجات منخفضة .

51) يقصد بثبات الاختبار ....
أ) قدرة الاختبار على قياس ما يراد قياسه .
ب)شمول أسئلة الاختبار لكل المقرر الذي درسه التلاميذ .
ج)الحصول على نتائج مشابهة عند إعادة تطبيق الاختبار على نفس المجموعة
د) اختلاف النتائج عند إعادة تطبيق الاختبار على نفس المجموعة

52) أي أنواع الصدق التالية يناسب بدرجة اكبر الاختبارات التحصيلية ؟ ....
أ) صدق التكوين الفرضي
ب) الصدق التلازمي
ج) الصدق التنبؤي
د)صدق المحتوى

53) أي الخصائص التالية تتأثر بدرجة اكبر بقلة عدد مفردات الاختبار ؟
أ) الموضوعية ب) التقنين ج)الثبات د)الصدق

54) يحتمل ان تنخفض قيمة معامل ثبات الاختبار إذا :
أ) زادت عدد مفرداته
ب)تشتت درجاته تشتتاً كبيراً
ج) تساوت قيم الوسط والوسيط والمنوال
د)اشتمل على مفردات لا تقيس ما وضع لقياسه

55) طبق اختبار مرتين في يومين متتاليين على مجموعة من الطلاب ، ووجد ان معامل الارتباط بين العلامات في المرتين (0.95) أي الاستنتاجات التالية صحيحة
أ) العلامات تتميز بدرجة عالية من الثبات
ب) العلامات تتميز بدرجة عالية من الصدق
ج)العلامات غير مستقرة إلى حد كبير
د)العلامات غير مميزة

56) إن تطبيق اختبار ما على مجموعة من الطلبة ذوي التحصيل المنخفض يؤدي غالباً الانتهاء بمعامل ثبات ......
أ) متوسط ب) منخفض ج) عالي د) لا يمكن التنبؤ

57) أهم خاصيتين للاختبار ككل هما......

أ) الصعوبة والتمييز

ب) الصدق والصعوبة

ج) الثبات والصدق

د) التمييز والثبات

58) يمكن ان يتبع المعلم جميع الطرق الاتية عند تصحيح أسئلة اختبارات المقال ما عدا :

أ) اعتماد ورقة إجابة الطالب المميز كنموذج للإجابة.

ب) الإعداد المسبق لنموذج تصحيح مناسب.

ج)استبعاد أسماء الطلاب من أوراق الإجابة قبل البدء في التصحيح .

د)تصحيح كل سؤال في جميع أوراق الإجابة قبل الانتقال إلى السؤال التالي.

59) إذا حصل طالب في اختبار على درجة خام تناظر المئيني 56 . فهذا يعني ان:

أ) أداءه يفوق 56 طالباً من أقرانه من المختبرين

ب) أداء 56% من الطلاب أفضل من أدائه في الإختبار

ج) أداءه يعادل أو أفضل من أداء 56%من أقرانه من المختبرين

د) درجة الطالب تزيد بقدر 6 درجات عن متوسط درجات الإختبار

60) فيما يلي المتوسط والانحراف المعياري لدرجات اختبارين :

	الوسط الحسابي	الانحراف المعياري
الإختبار الأول	40	5
الإختبار الثاني	34	2

فإذا كانت هذه العلامات تتوزع اعتدالياً ، وحصل طالب على الدرجتين 45 ، 37 في كل من الاختبارين على الترتيب . فأي مما يأتي يعد صحيحاً فيما يتعلق بأداء الطالب ؟

أ) أداؤه في الإختبار الأول افضل من أدائه في الإختبار الثاني

ب) أداؤه في الإختبار الثاني أفضل من أدائه في الإختبار الأول

ج) أداؤه متعادل في كل من الاختبارين

د) المعلومات المعطاة غير كافية للحكم على أداء الطالب

61) إذا حصل طالب على الدرجة الخام 20 في اختبار ما ، وكان متوسط درجات الصف 8 ، والانحراف المعياري 4 ، فإن درجتة المعيارية تساوي:

أ)-2 ب)-1 ج)+2 د)+3

62) أجاب طالب عن 25 سؤال إجابة صحيحة في اختبار رياضيات يشتمل على 50 سؤال . ما التفسير المحتمل لأداء الطالب ؟

أ) يحتاج الطالب لدروس تقوية في الرياضيات

ب) يعرف حوالي نصف مادة الرياضيات التي تضمنها الإختبار

ج) مركز الطالب النسبي هو المئيني 50

د) المعلومات المعطاة غير كافية للتفسير

63) يتطلب إعطاء الدرجة على أساس تقدير(محكي المرجع) :

أ) مقارنة التلميذ بنفسه

ب)مقارنة التلميذ بزملائه

ج)مقارنة التلميذ بنسبة مئوية مطلقة

د) جميع ما سبق

64) توزيع درجات اختبار تحصيلي في الرياضيات طبق على تلاميذ الصف الثاني لكنه مصمم لتلاميذ الصف الخامس يحتمل بدرجة اكبر ان يكون ...

أ) ثنائي المنوال

ب)ملتوياً التواء موجباً

ج) ملتوياً التواء سالباً

د) اعتداليا ً

65) إذا كان الانحراف المعياري لعينة من العلامات صغيراً ، فيمكن القول بأن تلك العلامات ..

أ)متغيرة ب)ثابتة ج)متطرفة د) متجانسة

66)قيمة الوسط الحسابي للعلامات 1، 2 ، 4 ، 6 ،8، 9 هو....

أ)6 ب)30 ج) 5 د)4

67)إذا كان مجموع مربعات انحرافات (5) قيم عن وسطها الحسابي هو (20) فأن الانحراف المعياري لها = ...

أ) 2　　ب) 4　　ج) 1　　د)16

68) طول الفئة 5 – 11 هو .....

أ) 5　　ب) 6　　ج) 7　　د) 8

69) إذا تم تبويب العلامات 17 ، 6 ، 8 ،10، 3 ، 15 ، 14 ، 6 ،8، 9 في جدول توزيع تكراري فأن عدد العلامات التي تقع في الفئة التي طولها 6 ومنتصفها 6.5 هو ...

أ)5　　ب) 6　　ج) 4　　د) 3

70) في اختبار تحصيلي لأحد الصفوف وجد ان قيمة الوسط الحسابي 40 وقيمة المنوال 70 وقيمة الوسيط 50 فأن هذا الاختبار يمكن ان يوصف بأنه

أ)صعب جداً.

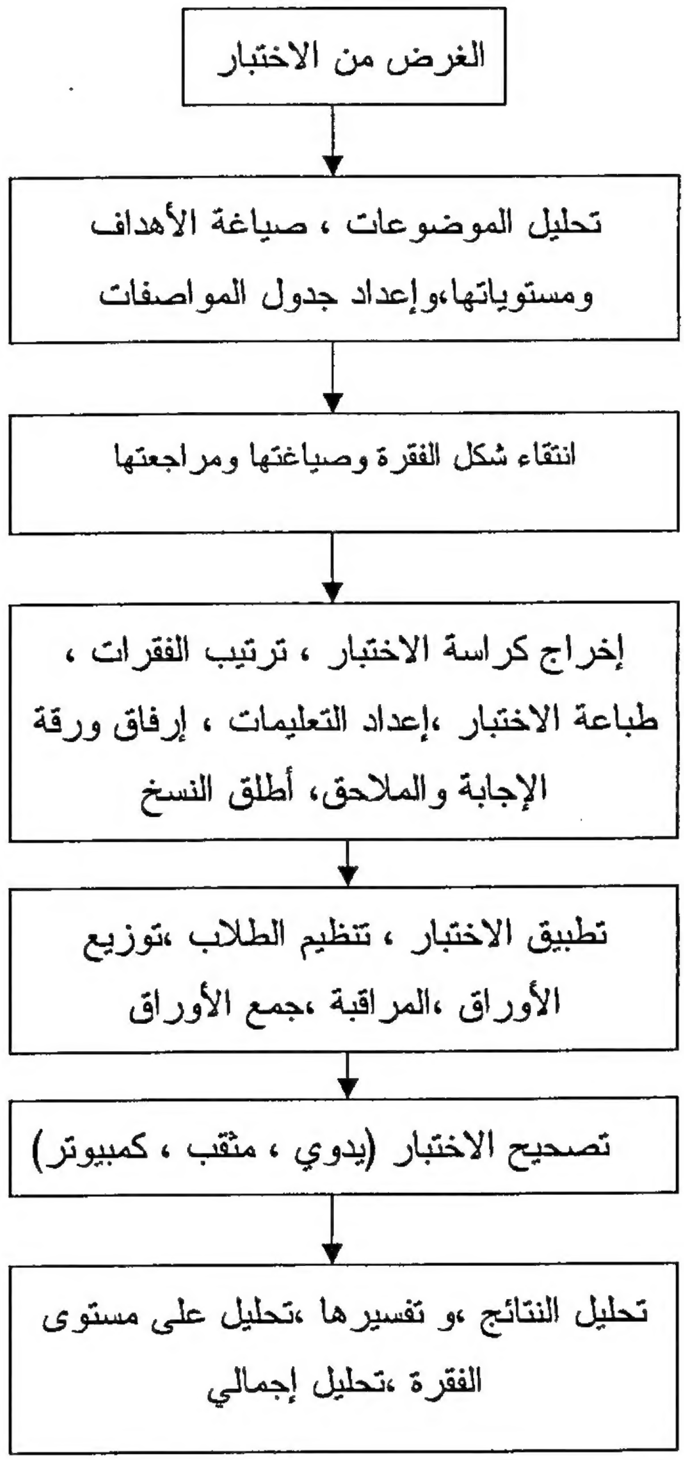
ب)سهل جداً.

ج)متوسط الصعوبة.

د) لا يمكن التنبؤ من هذه المعطيات.

رقم الفقرة	البدائل			
	أ	ب	ج	د
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				

رقم الفقرة	البدائل			
	أ	ب	ج	د
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				


الغرض من الاختبار
تحليل الموضوعات ، صياغة الأهداف ومستوياتها،وإعداد جدول المواصفات
انتقاء شكل الفقرة وصياغتها ومراجعتها
إخراج كراسة الاختبار ، ترتيب الفقرات ، طباعة الاختبار ،إعداد التعليمات ، إرفاق ورقة الإجابة والملاحق، أطلق النسخ
تطبيق الاختبار ، تنظيم الطلاب ،توزيع الأوراق ،المراقبة ،جمع الأوراق
تصحيح الاختبار (يدوي ، مثقب ، كمبيوتر)
تحليل النتائج ،و تفسيرها ،تحليل على مستوى الفقرة ،تحليل إجمالي